

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:02:27 ; Search time 42 Seconds  
(without alignments)  
600.974 Million cell updates/sec

Title: US-09-634-109D-2  
Perfect score: 1634  
Sequence: 1 MGRWVNSYDGEFFLGIFG.....GEVWGALRKGLDRCRIGSQH 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 segs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCR\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/PCR\_PUBCOMB.pep:\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1432	87.6	315	10	US-09-886-055-175 Sequence 175, App
2	880.5	53.9	369	10	US-09-886-055-457 Sequence 457, App
3	879.5	53.8	316	10	US-09-886-055-477 Sequence 477, App
4	869	53.2	316	10	US-09-826-508-24 Sequence 24, Appl
5	858	52.5	317	10	US-09-886-055-429 Sequence 429, App
6	846	51.8	314	10	US-09-886-055-479 Sequence 479, App
7	815	49.9	314	10	US-09-761-288-16 Sequence 16, Appl
8	815	49.9	314	10	US-09-886-055-143 Sequence 143, App
9	815	49.9	317	10	US-09-761-288-18 Sequence 18, Appl
10	814	49.8	312	10	US-09-886-055-421 Sequence 421, App
11	813.5	49.8	348	10	US-09-886-055-433 Sequence 433, App
12	810	49.6	305	10	US-09-761-288-73 Sequence 73, Appl
13	810	49.6	305	10	US-09-761-288-75 Sequence 75, Appl
14	810	49.6	305	10	US-09-761-288-79 Sequence 79, Appl
15	810	49.6	315	10	US-09-886-055-435 Sequence 435, App
16	798	48.8	338	10	US-09-886-055-427 Sequence 427, App
17	781	47.8	312	10	US-09-886-055-473 Sequence 473, App
18	762	46.6	331	10	US-09-886-055-475 Sequence 475, App
19	746.5	45.7	323	10	US-09-886-055-163 Sequence 163, App

20	741.5	45.4	348	10	US-09-886-055-165 Sequence 165, App
21	708	43.3	310	10	US-09-747-835A-63 Sequence 63, Appl
22	702	43.0	353	10	US-09-747-835A-47 Sequence 47, Appl
23	700	42.8	311	10	US-09-864-761-46639 Sequence 46639, A
24	692	42.4	312	9	US-09-974-591-16 Sequence 16, Appl
25	692	42.4	312	9	US-09-974-591-18 Sequence 18, Appl
26	691	42.3	310	9	US-09-800-321A-36 Sequence 36, Appl
27	691	42.3	316	10	US-09-886-055-179 Sequence 179, App
28	691	42.3	357	9	US-09-800-321A-37 Sequence 37, Appl
29	690	42.2	314	10	US-09-771-730-32 Sequence 32, Appl
30	690	42.2	357	9	US-09-800-321A-35 Sequence 35, Appl
31	689	42.2	307	10	US-09-771-730-112 Sequence 112, App
32	689	42.2	314	10	US-09-771-730-24 Sequence 24, Appl
33	689	42.2	319	10	US-09-886-055-387 Sequence 387, App
34	689	42.2	330	10	US-09-886-055-199 Sequence 199, App
35	688	42.1	307	10	US-09-771-730-106 Sequence 106, App
36	688	42.1	314	10	US-09-771-730-22 Sequence 22, Appl
37	688	42.1	314	10	US-09-771-730-36 Sequence 36, Appl
38	687	42.0	305	10	US-09-771-730-114 Sequence 114, App
39	687	42.0	305	10	US-09-771-730-144 Sequence 144, App
40	687	42.0	314	10	US-09-771-730-26 Sequence 26, Appl
41	687	42.0	314	10	US-09-771-730-34 Sequence 34, Appl
42	686	42.0	304	10	US-09-771-730-136 Sequence 136, App
43	685	41.9	307	10	US-09-771-730-107 Sequence 107, App
44	685	41.9	307	10	US-09-771-730-113 Sequence 113, App
45	685	41.9	317	10	US-09-771-730-108 Sequence 108, App

ALIGNMENTS

RESULT 1  
US-09-886-055-175  
Sequence/175, Application US/09886055  
Patent No. US20020132273A1  
GENERAL INFORMATION:  
APPLICANT: STRYER, LUBERT  
APPLICANT: ZOZULYA, SERGEY  
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
FILE REFERENCE: 078003-0277150  
CURRENT APPLICATION NUMBER: US/09/886,055  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/213,812  
PRIOR FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 522  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 175  
LENGTH: 315  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-886-055-175

Query Match	87.6%	Score 1432;	DB 10;	Length 315;
Best Local Similarity	87.0%	Pred. No. 5.3e-126;		
Matches 274;	Conservative 13;	Mismatches 28;	Indels 0;	Gaps 0;
QY	1	MGRWVNSYDGEFFLGIFSHSQTDLVFSAVMVFTVALCGNVLLIFLIYDAGLHTPM	60	
Db	1	METWVNSYDGEFFLGIFSHSTADLVFSAVMVFTVALCGNVLLIFLIYDAPHLTPM	60	
QY	61	YFELSLSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQGFVSLVSGSLGLGMA	120	
Db	61	YFELSLSLMDLMLVCTNVPKMAANFLSGRKSISFVGCIGIQGFVSLVSGSLGLGMA	120	
QY	121	YDHYVAVSHPLHYPTILMNQRYCLQITGSSMAFGIIDGVIONVAAAGLPYCGSRVDHFFC	180	
Db	121	YDRYVAISHPLHYPTILMNQRYCLQITGSSMAFGIIDGVIONVAAAGLPYCGSRVDHFFC	180	
QY	181	EVQALLKLACADTSLFDTLLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA	240	
Db	181	EMLSLKLACADTSLFKEKVFACCVFMLLPFSIIVASYAHILGTVLQMHSAQAWKKALA	240	



QY	186	LKLACADTSLFDTLLFACCVFMLLPESIIIMASYACILGAVLRIRSAQWKKLATCSSH	245
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	185	LILSCNDSTFEKVI FICSI VMLVFPAII IASVAGVILAVIHNGSGEGRKRAFTTCSH	244
QY	246	LTAVTLFYGAAMFYLRPRRYRAPSHDKVASIFETVTLPLMPLNPLYSLRNGEYWGALRK	304
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	245	LMVVGMEYAGLEMYIQPTSDRSPTODKLVSEYFTLLTPLMNLPLYSLRNKEVTRAFMK	303

RESULT 5  
US-09-886-055-429  
; Sequence 429, Application US/09886055

```

; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 429
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-429

```

Query Match	52.5%;	Score 858;	DB 10;	Length 324;
Best Local Similarity	53.9%;	Pred. No. 2e-72;		
Matches 160;	Conservative 51;	Mismatches 86;	Indels 0;	Gaps 0;

[illegible]

RESULT 6  
US-09-886-055-479  
; Sequence 479, Application US/09886055.  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886, 055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213, 812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 479
; LENGTH: 317
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-886-055-479

```

Query Match	51.88;	Score 846;	DB 10;	Length 317;
Best Local Similarity	50.7%;	Pred. No. 2.6e-71;		
Matches 154;	Conservative 60;	Mismatches 90;	Indels 0;	Gaps 0;

QY	6	NOSYTDFEFLIGIFSHSQTDVLVFSANWVFVALCGNVLLIFELLYDAGLHFPMPYFFLS	65
Dd	5	NYSVYADFTILGLFNSNARPPMLLFALLIVFLTSTASNVKILIHIDSRHLHPMYFFLS	64
QY	66	QLSLMDLMVCNIVPKMAANFLSGRKSISFVGCGIQIGFEVSLVGSGLLGIMAYDHV	125
Dd	65	QLSLRDILEYSTIVPKMLVDQVMQSORASIFAGCTAQHFLLYLTLGAEFFLLGLMSYDRYV	124
QY	126	AVSHPLHYPLMNQRVCLOITGSSWAEGIIDGVIOVAAAMGLPYCGRSRVDPHEFCVOAL	185
Dd	125	AICNPPLHYPLMSRKICMULIVAAMWGCSIDGFLLPVTMQFPCASREINHFECEVPAL	184
QY	186	LKLACADTSLFDTLFACCVEFMLLPESIIMASYACILGAVLRIRSAQAWMKALATCSSH	245
Dd	185	LKLSCDTDSAYETAMYVCCIMMLLIPPSVISGSYRILITVRNSEAEGRGKAVATCSSH	244
QY	246	LTAVTLFYGAMFMYLRRRYRAPASHDKVASIFYTVLTPMLNPLIYSLRNGEVMGALRKG	305
Dd	245	MVVVSLEFYGAMTYTVLEPHSYHTPEQDKAVSAFYITLTPMLNPLIYSLRNKDVTGALQKV	304
QY	306	LDRC	309
Dd	305	VGRC	308

RESULT 7  
US-09-761-288-16  
; Sequence 16, Application US/09761288  
; Patent No. US20020065405A1

```

1  APPLICANT: Padigaru, Muralidhara
2  APPLICANT: Prayaga, Sudhirdas
3  APPLICANT: Taupier, Raymond J
4  APPLICANT: Mishra, Vishnu
5  APPLICANT: Tchernev, Velizar
6  APPLICANT: Spytek, Kimberly
7  APPLICANT: Li, Li
8  TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding
9  FILE REFERENCE: 15966-638
10 CURRENT APPLICATION NUMBER: US/09/761,288

```

```
?
?      PRIOR APPLICATION NUMBER: 60/177,839
?      PRIOR FILING DATE: 2000-01-25
?      PRIOR APPLICATION NUMBER: 60/176,134
?      PRIOR FILING DATE: 2000-01-14
?      PRIOR APPLICATION NUMBER: 60/175,989
?      PRIOR FILING DATE: 2000-01-13
?      PRIOR APPLICATION NUMBER: 60/218,324
?      PRIOR FILING DATE: 2000-07-14
?      PRIOR APPLICATION NUMBER: 60/220,253
?      PRIOR FILING DATE: 2000-07-24
?      PRIOR APPLICATION NUMBER: 60/178,191
?      PRIOR FILING DATE: 2000-01-26
?      PRIOR APPLICATION NUMBER: 60/178,227
?      PRIOR FILING DATE: 2000-01-26
?      PRIOR APPLICATION NUMBER: 60/220,590
?      PRIOR FILING DATE: 2000-07-25
?      NUMBER OF SEQ ID NOS: 95
?      SOFTWARE: Patentin Ver. 2.0
?      SEQ ID NO 16
?      LENGTH: 314
?      TYPE: prt
?      ORGANISM: Homo sapiens
```

US-09-761-288-16

Query Match 49.9%; Score 815; DB 10; Length 314;  
Best Local Similarity 54.1%; Pred. No. 2e-68;  
Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNQSYTDGFFLLGIFSHSQTDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYFLL 64  
Db 4 VNQSVASDFILVGLFSGSRQLLFSLVAVMEVIGLLGNVLLFLIRVDSRLHTPMYFLL 63  
QY 65 SOLSLMDLMLVCNIVPKMANFLSGRKSISFVCGCIQIGFEVSLVG-SEGILLGLMAYDH 123  
Db 64 SOLSLFDIGCPMWTIPKMASDFLRGEGATSYCGGAQI-FFLLTMGVAEGVLLVMSYDR 122  
QY 124 YVAVSHPLHYFILMNQRCVCIQTGSSWAFGIIDGVIQNVAAAGLPGCGSRVDHFFCEVQ 183  
Db 123 YVAVCQPLQYVLMRQVCLLMGSSWVGVNLNASIQTSITLHPYCASRIVDHFFCEVP 182  
QY 184 ALLKLACADTSLFDTLFACCVFMLLPFSITIMASYACILGAVLRIRSAQAMKKALATCS 243  
Db 183 ALLKLSCADTCAYEMALSTSGVLILMLPLSLIATSYGHVLAQAVLSMRSEARHKAVTTC 242  
QY 244 SHLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLMNLPLIYSLRNGEVMGALR 303  
Db 243 SHITVGLFYGAAVFMYMPCAYHSPQDQNVSLFYSLVTPPLNPLIYSLRNPVMMALV 302  
QY 304 KGLDR 308  
Db 303 KVLRSR 307

RESULT 8  
US-09-886-055-143

; Sequence 143, Application US/09886055  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886, 055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213, 812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 143  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-886-055-143

Query Match 49.9%; Score 815; DB 10; Length 314;  
Best Local Similarity 54.1%; Pred. No. 2e-68;  
Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNQSYTDGFFLLGIFSHSQTDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYFLL 64  
Db 4 VNQSVASDFILVGLFSGSRQLLFSLVAVMEVIGLLGNVLLFLIRVDSRLHTPMYFLL 63  
QY 65 SOLSLMDLMLVCNIVPKMANFLSGRKSISFVCGCIQIGFEVSLVG-SEGILLGLMAYDH 123  
Db 64 SOLSLFDIGCPMWTIPKMASDFLRGEGATSYCGGAQI-FFLLTMGVAEGVLLVMSYDR 122  
QY 124 YVAVSHPLHYFILMNQRCVCIQTGSSWAFGIIDGVIQNVAAAGLPGCGSRVDHFFCEVQ 183  
Db 123 YVAVCQPLQYVLMRQVCLLMGSSWVGVNLNASIQTSITLHPYCASRIVDHFFCEVP 182  
QY 184 ALLKLACADTSLFDTLFACCVFMLLPFSITIMASYACILGAVLRIRSAQAMKKALATCS 243  
Db 183 ALLKLSCADTCAYEMALSTSGVLILMLPLSLIATSYGHVLAQAVLSMRSEARHKAVTTC 242

QY 244 SHLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLMNLPLIYSLRNGEVMGALR 303  
Db 243 SHITVGLFYGAAVFMYMPCAYHSPQDQNVSLFYSLVTPPLNPLIYSLRNPVMMALV 302  
QY 304 KGLDR 308  
Db 303 KVLRSR 307

RESULT 9

US-09-761-288-18  
; Sequence 18, Application US/09761288  
; Patent No. US20020065405A1  
; GENERAL INFORMATION:

; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Prayaga, Sudhirdas  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Li, Li  
; TITLE OF INVENTION: No. US20020065405A1e1 Polypeptides and Nucleic Acids Encoding  
; FILE REFERENCE: 15966-638  
; CURRENT APPLICATION NUMBER: US/09/761, 288  
; CURRENT FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/177, 839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176, 134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175, 989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218, 324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220, 253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178, 191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178, 227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220, 590  
; PRIOR FILING DATE: 2000-07-25  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-288-18

Query Match 49.9%; Score 815; DB 10; Length 317;  
Best Local Similarity 54.1%; Pred. No. 2e-68;  
Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNQSYTDGFFLLGIFSHSQTDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYFLL 64  
Db 4 VNQSVASDFILVGLFSGSRQLLFSLVAVMEVIGLLGNVLLFLIRVDSRLHTPMYFLL 63  
QY 65 SOLSLMDLMLVCNIVPKMANFLSGRKSISFVCGCIQIGFEVSLVG-SEGILLGLMAYDH 123  
Db 64 SOLSLFDIGCPMWTIPKMASDFLRGEGATSYCGGAQI-FFLLTMGVAEGVLLVMSYDR 122  
QY 124 YVAVSHPLHYFILMNQRCVCIQTGSSWAFGIIDGVIQNVAAAGLPGCGSRVDHFFCEVQ 183  
Db 123 YVAVCQPLQYVLMRQVCLLMGSSWVGVNLNASIQTSITLHPYCASRIVDHFFCEVP 182  
QY 184 ALLKLACADTSLFDTLFACCVFMLLPFSITIMASYACILGAVLRIRSAQAMKKALATCS 243  
Db 183 ALLKLSCADTCAYEMALSTSGVLILMLPLSLIATSYGHVLAQAVLSMRSEARHKAVTTC 242  
QY 244 SHLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLMNLPLIYSLRNGEVMGALR 303  
Db 243 SHITVGLFYGAAVFMYMPCAYHSPQDQNVSLFYSLVTPPLNPLIYSLRNPVMMALV 302



OY 304 KGLDR 308  
| | |  
Db 303 KVLNR 307

## RESULT 10

US-09-886-055-421  
; Sequence 421, Application US/09886055  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886,055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213,812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 421  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-886-055-421

Query Match 49.8%; Score 814; DB 10; Length 312;

Best Local Similarity 52.9%; Pred. No. 2.4e-68;

Matches 163; Conservative 50; Mismatches 91; Indels 4; Gaps 4;

OY 6 NOSYTDGFFLLGIFSHSQTDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYFPLS 65  
| | : | | | | | | | | : | : | | | : | | | | | | | | | | | | | |  
Db 5 NOTSTD-FILLGLFPPIIDLFFELIYFIEMALIGNLSMILLIFLIDTHLHTPMYFPLS 63  
OY 66 QLSIMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQGFVSLVSGELLLGLMAYDHYV 125  
| | | | | : | | | | | : | | | | | | | | | | | | | | | | | | | |  
Db 64 QLSLIDINYSTIVPKMASDFLHGKNSISFTGCGIQSFFFLALGAEALLASMAVDRI 123  
OY 126 AVSHPLHYPIILMNQRYC-LQITGSSWAFGIIDGVIQVNAAMGLPYCGSRSDHFFCEYQA 184  
| : | | | | | : | | | | | | | | | | : | : | | | : | | | | | | | |  
Db 124 AICFPLHYLRMSKRYCLMITG-SWIGSINACAHFYVVLHPIYCRRAINHFCDVPA 182  
OY 185 LKLIACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALATCSS 244  
: | | | | : : : | : | | | | | | | | | | : | | | | | | | | | |  
Db 183 MVTIACMDTWVEGTVEFSATIFLVFEPFISCSYGVLFVAVYHMKSAEGRKKAAYLTCSY 242  
OY 245 HLTAVTLFYGAAMFYVLRPRRYRAPSHDKVASIFVTVLTPMLNPLIYSLRNGEVMGALRK 304  
| | | | : | : | | | | : | : | | | | | | | | | | | | | | | | | | |  
Db 243 HLTAVTFYAPFVTVLRPRSLRSPTEBKVLAFFYTLTPMLNPLIYSLRNKEVMGALTR 302  
OY 305 GLDR-CRI 311  
| | :  
Db 303 VSQRICSV 310

## RESULT 11

US-09-886-055-433  
; Sequence 433, Application US/09886055  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886,055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213,812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 433

; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-886-055-433

## Query Match 49.8%; Score 813.5; DB 10; Length 348;

Best Local Similarity 52.0%; Pred. No. 3.1e-68;

Matches 158; Conservative 47; Mismatches 98; Indels 1; Gaps 1;

OY 4 WV-NQSYTDGFFLLGIFSHSQTDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYF 62  
| : | : | | | | | : | : | : | | | : | | | | | | | | | | | | | |  
Db 34 WMANHTGMSDFILLGLFRQSKHPALCVIVFVFLMALSGNAVLLILHCDALHTPMYF 93  
OY 63 FLSQSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQGFVSLVSGELLLGLMAYD 122  
| : | | | | | : | | | : | | | | | | | | | | | | | | | | | | | |  
Db 94 FISQSLMDMAYISVTVPKMLLDQVMGVNKRISAPECGMQMFEFVTLAGSEFFELATMAYD 153  
OY 123 HYAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIQVNAAMGLPYCGSRSDHFFCEY 182  
| | : | | | | | | | | : | | | | | | | | | | | | | | | | | | | |  
Db 154 RYVALCHPLRYPVLMNHRVCLFLSSGCVFLGSDGFTPTPTMTFPPFRGSRREIHHFCEY 213  
OY 183 QALLKLIACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALATC 242  
| : | | : | | | | : : | | | | | | | | | | | | | | | | | | | | | |  
Db 214 PAVLNLSGSDTSLYEIRFMYLCVLMILIPVLISSSYLLILFTIHGMNSAEGRKKAFAATC 273  
OY 243 SSSLTAVTLFYGAAMFYVLRPRRYRAPSHDKVASIFVTVLTPMLNPLIYSLRNGEVMGAL 302  
| | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | |  
Db 274 SSSLTAVTLFYGAAYTYMLPSSYHFPBKDMVSVFYTLTPVNPPLIYSLRNKDVMGAL 333  
OY 303 RKGL 306  
: | |  
Db 334 KKMFL 337

## RESULT 12

US-09-761-288-73  
; Sequence 73, Application US/09761288  
; Patent No. US20020065405A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Prayaga, Sudhirdas  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Li, Li  
; TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding  
; FILE REFERENCE: 15966-638  
; CURRENT APPLICATION NUMBER: US/09/761,288  
; CURRENT FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/177,839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176,134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175,989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218,324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178,191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178,227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220,590  
; PRIOR FILING DATE: 2000-07-25  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 73  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-761-288-73

[illegible]

```

RESULT 13
US-09-761-288-75
Sequence 75, Application US/09761288
Patent No. US20020065405A1
GENERAL INFORMATION:
APPLICANT: Padigaru, Muralidhara
APPLICANT: Prayaga, Sudhirdas
APPLICANT: Taupier, Raymond J
APPLICANT: Mishra, Vishnu
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Li, Li
TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam
FILE REFERENCE: 15966-638
CURRENT APPLICATION NUMBER: US/09/761,288
CURRENT FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/177,839
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 60/176,134
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: 60/175,989
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: 60/218,324
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/220,253
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: 60/178,191
PRIOR FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: 60/178,227
PRIOR FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: 60/220,590
PRIOR FILING DATE: 2000-07-25
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 75
LENGTH: 305
TYPE: PRT
ORGANISM: Homo sapiens
US-09-761-288-75

```

Query Match	49.6%;	Score 810;	DB 10;	Length 305;
Best Local Similarity	54.1%;	Pred. No. 5.6e-68;		
Matches 164;	Conservative 48;	Mismatches 89;	Indels 2;	Gaps 2

[illegible]

```

RESULT 14
; US-09-761-288-79
; Sequence 79, Application US/09761288
; Patent No. US20020065405A1
;
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Prayaga, Sudhirdas
; APPLICANT: Taupier, Raymond J
; APPLICANT: Mishra, Vishnu
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Li, Li
; TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding
; FILE REFERENCE: 15966-638
; CURRENT APPLICATION NUMBER: US/09/761,288
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/177,839
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/176,134
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/175,989
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/218,324
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,253
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/178,191
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/178,227
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/220,590
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-761-288-79

```

Query Match	49.68;	Score 810;	DB 10;	Length 305;
Best Local Similarity	54.1%;	Pred. No. 5.6e-68;		
Matches 164;	Conservative 48;	Mismatches 89;	Indels 2;	Gaps 2;
QY	5 VNOSYTDGFELLGIFSHSQTDLVLFSAVMVFVTALCGNVLLIFLIYLDAGLHTPMYFEL	64		
Db	4 VNOSVASDFILVGLFSHSGSROLLFSLVAVMFVIGLLGNTVLLFLLIRDSRLHTPMYFEL	63		

[illegible]

Search completed: April 28, 2003, 14:07:52  
Job time : 43 secs

```

RESULT 15
US-09-886-055-435
; Sequence 435, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 435
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-435

```

Query Match	49.68;	Score 810;	DB 10;	Length 315;
Best Local Similarity	51.38;	Pred. No. 5.8e-68;		
Matches 156; Conservative	48;	Mismatches 100;	Indels 0;	Gaps 0

[illegible]

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:52:36 ; Search time 39 Seconds  
(without alignments)  
1076.255 Million cell updates/sec

Title: US-09-634-109D-2

Perfect score: 1634  
Sequence: 1 MGRWVNQSYTDGFLLGIFS.....GEVWGALRKGLDRCRIGSQH 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1634	100.0	315	22	AAU00529	Human CON167 G pro
2	1626	99.5	326	22	AAG71793	Human Olfactory re
3	1620	99.1	315	23	ABG68141	Human G-protein co
4	1432	87.6	315	22	AAG71699	Human Olfactory re
5	1432	87.6	315	22	AAG71792	Human Olfactory re
6	1432	87.6	315	22	AAG72461	Human OR-like poly
7	1432	87.6	315	23	AAU95619	Human Olfactory an
8	1432	87.6	315	23	AAE18656	Human G-protein co
9	1432	87.6	315	23	AAU85223	G-coupled olfactor
10	1360	83.2	271	23	AAE16175	Human G-protein co

11	1150	70.4	252	22	AAU24603	Human olfactory re
12	1146	70.1	263	23	ABP61132	Human GPCR7 protei
13	1044	63.9	223	21	AAU96666	Murine olfactory r
14	1044	63.9	223	21	AAU96670	Murine olfactory r
15	1040	63.6	223	21	AAU96668	Murine olfactory r
16	911	55.8	320	23	AAU95784	Human olfactory an
17	902	55.2	311	23	AAU95780	Human olfactory an
18	896	54.8	320	23	AAU95782	Human olfactory an
19	894	54.7	312	23	AAU95781	Human olfactory an
20	882.5	54.0	316	23	ABP81453	Human GPCR7 prote
21	880.5	53.9	311	22	AAG71713	Human olfactory re
22	880.5	53.9	318	23	AAU95679	Human olfactory an
23	880.5	53.9	369	22	AAU24744	Human olfactory re
24	880.5	53.9	369	23	AAU85364	G-coupled olfactor
25	879.5	53.8	316	22	AAU24754	Human olfactory re
26	879.5	53.8	316	23	ABP81452	Human GPCR7a prote
27	879.5	53.8	316	23	AAU85374	G-coupled olfactor
28	878	53.7	312	23	AAU95783	Human olfactory an
29	870	53.2	289	23	AAU95669	Human olfactory an
30	869	53.2	316	19	AAW75960	Human olfactory OL
31	862	52.8	321	22	AAG71467	Human olfactory re
32	858	52.5	324	22	AAU24730	Human olfactory re
33	858	52.5	324	23	AAU95680	Human G-protein co
34	858	52.5	324	23	AAE18655	Human G-coupled olfactor
35	858	52.5	324	23	AAU85350	Human GPCR4b prote
36	849	52.0	317	23	ABP81447	Human olfactory re
37	846	51.8	317	22	AAU24755	Human GPCR4a prote
38	846	51.8	317	23	ABP81446	G-coupled olfactor
39	846	51.8	317	23	AAU85375	G-protein coupled
40	843	51.6	318	22	AAE10683	Human GPCR8a prote
41	840	51.4	317	23	ABP81455	Human GPCR8b prote
42	840	51.4	317	23	ABP81456	Human G protein co
43	836	51.2	317	23	ABP51586	Human G-protein co
44	828	50.7	312	23	AAE18265	Human GPCR7c prote
45	826	50.6	275	23	ABP81454	

ALIGNMENTS

*Sam*

RESULT 1	AAU00529	standard; Protein; 315 AA.
ID	AAU00529	
AC	AAU00529;	
XX	14-MAY-2001	(first entry)
DT		
XX		
DE	Human CON167 G protein seven transmembrane receptor.	
XX		
KW	CON167; G protein coupled receptor; transmembrane receptor;	
KW	neurological disorder; psychiatric disease; schizophrenia; depression;	
KW	anxiety; bipolar disorder; affective disorder; epilepsy; neuritis;	
KW	attention deficit hyperactivity disorder; ADHD; neuroasthenia; neuropathy;	
KW	neurosis; Alzheimer's disease; Parkinson's disease; migraine; therapy;	
KW	senile dementia; hybridisation assay; diagnosis; transgenic animal.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	
FT	Domain	
FT	Location/Qualifiers	
FT	/note= "Transmembrane-spanning domain"	
FT	60..79	
FT	/note= "Transmembrane-spanning-domain"	
FT	92..121	
FT	/note= "Transmembrane-spanning-domain"	
FT	151..170	
FT	/note= "Transmembrane-spanning-domain"	
FT	196..220	
FT	/note= "Transmembrane-spanning-domain"	
FT	242..260	
FT	/note= "Transmembrane-spanning-domain"	
FT	274..294	
FT	Domain	

FT /note= "Transmembrane spanning-domain"  
XX  
PN WO200114554-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 08-AUG-2000; 2000WO-US21566.  
XX  
PR 19-AUG-1999; 99US-0377563.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Vogel I G, Wood LS;  
XX  
DR WPI: 2001-218450/22.  
DR N-PSDB; AAS00555.  
XX  
PT Novel purified isolated seven transmembrane receptor polypeptide  
PT (G-protein coupled receptor) useful for treating neurological and  
PT psychiatric diseases such as schizophrenia, depression, anxiety,  
PT bipolar disease and affective disorder  
XX  
PS Claim 1; Page 63-64; 72pp; English.  
XX  
CC The sequence represents human CON167 G protein seven transmembrane  
CC receptor. The protein and its corresponding DNA are useful for  
CC modulating activity of CON167 in a mammal comprising cells that express  
CC CON167, preferably in a human suffering from a neurological disorder  
CC and/or psychiatric diseases such as schizophrenia, depression, anxiety,  
CC bipolar disease, affective disorder, attention deficit hyperactivity  
CC disorder (ADHD), epilepsy, neuritis, neuroasthenia, neuropathy, neuroses,  
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
CC The polynucleotides are useful in hybridisation assays to detect the  
CC capacity of cells to express CON167, for large scale expression of CON167  
CC required for research, for identification and isolation of new  
CC polynucleotides encoding the related CON167 polypeptides, in diagnostic  
CC methods for identifying any genetic alteration in a CON167 locus that  
CC underlies a disease state (which is useful for selection of therapeutic  
CC strategies), and for the development of transgenic animals that fall to  
CC express functional CON167 or that express a variant of CON167.  
XX  
SQ Sequence 315 AA;  
Query Match 100.0%; Score 1634; DB 22; Length 315;  
Best Local Similarity 100.0%; Pred. No. 9.8e-179;  
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60  
Db 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60  
QY 61 YFFLSQLSLMDLMLVCNIVPRMAANFLSGRKSISFVGCGIQIGFFVSLVSGEGLLGLMA 120  
Db 61 YFFLSQLSLMDLMLVCNIVPRMAANFLSGRKSISFVGCGIQIGFFVSLVSGEGLLGLMA 120  
QY 121 YDHYVAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
Db 121 YDHYVAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
QY 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
Db 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
QY 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMG 300  
Db 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMG 300  
QY 301 ALRKGIDRCRIGSQH 315  
Db 301 ALRKGIDRCRIGSQH 315  
RESULT 2

AAG71793  
ID AAG71793 standard; Protein; 326 AA.  
XX  
AC AAG71793;  
XX  
DT 31-JUL-2001 (first entry)  
XX  
DE Human olfactory receptor polypeptide, SEQ ID NO: 1474.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation.  
XX  
OS Homo sapiens.  
XX  
PN WO200127158-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 06-OCT-2000; 2000WO-US27582.  
XX  
PR 08-OCT-1999; 99US-0158615.  
XX  
PR 24-FEB-2000; 2000US-0184809.  
XX  
PA (DIGI-) DIGISCENTS.  
XX  
PI (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX  
DR WPI: 2001-290713/30.  
XX  
PT New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists  
XX  
PS Claim 11; Page 932-933; 1857pp; English.  
XX  
CC The present sequence is an olfactory receptor which is encoded by  
CC one of a number of novel polynucleotides. The polynucleotides can be  
CC used in screening for olfactory agonists and antagonists. The methods  
CC allow for the determination of primary scents and the identification  
CC of the odour receptors used to detect these primary scents. The methods  
CC also enable determination of secondary scents and the identification of  
CC combinations of odour receptors that are involved in detecting such  
CC secondary scents. This enables the construction of a scent representation  
CC (also called a scent fingerprint or scent profile), which may be used to  
CC re-create and edit scents. Libraries of olfactory receptors are useful  
CC for determining the interaction pattern of a composition with the  
CC receptors, and can be used for determining differences in the olfactory  
CC faculties of different individuals.  
XX  
SQ Sequence 326 AA;  
Query Match 99.5%; Score 1626; DB 22; Length 326;  
Best Local Similarity 99.7%; Pred. No. 8.5e-178;  
Matches 314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60  
Db 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60  
QY 61 YFFLSQLSLMDLMLVCNIVPRMAANFLSGRKSISFVGCGIQIGFFVSLVSGEGLLGLMA 120  
Db 61 YFFLSQLSLMDLMLVCNIVPRMAANFLSGRKSISFVGCGIQIGFFVSLVSGEGLLGLMA 120  
QY 121 YDHYVAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
Db 121 YDHYVAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
QY 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
Db 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
QY 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMG 300



|||||  
Db 241 TCSSHLTAVTLEFYGAAMFYLPRRYRAPSHDKVASIFYVTLTPLMPLIYSLRNGEVMG 300  
QY 301 ALRKGLDRCRIGSQH 315  
Db 301 ALRKGLDRCRIGSQH 315  
RESULT 3  
ABG68141  
ID ABG68141 standard; Protein; 315 AA.  
XX  
AC ABG68141;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GPCR) GPCR protein #8.  
XX  
KW G-protein coupled receptor; receptor; GPCR; GPCR; cardiomyopathy;  
KW atherosclerosis; diabetes; cell signal processing; cancer; trauma;  
KW metabolic pathway modulation; neuro-olfactory system; surgery;  
KW neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;  
KW uterus cancer; immune response; acquired immunodeficiency syndrome;  
KW AIDS; asthma; Crohn's disease; multiple sclerosis;  
KW Albright hereditary osteodystrophy.  
XX  
OS Homo sapiens.  
XX  
PN WO200250117-A2.  
XX  
PD 27-JUN-2002.  
XX  
PF 18-DEC-2001; 2001WO-US49077.  
XX  
PR 18-DEC-2000; 2000US-256635P.  
PR 21-DEC-2000; 2000US-257876P.  
PR 04-JAN-2001; 2001US-259743P.  
PR 10-JAN-2001; 2001US-260718P.  
PR 12-JAN-2001; 2001US-261498P.  
PR 24-JAN-2001; 2001US-263689P.  
PR 08-FEB-2001; 2001US-267464P.  
PR 22-FEB-2001; 2001US-271021P.  
PR 14-MAR-2001; 2001US-275946P.  
PR 23-MAR-2001; 2001US-278150P.  
PR 18-APR-2001; 2001US-284591P.  
PR 23-APR-2001; 2001US-285718P.  
PR 19-JUN-2001; 2001US-299327P.  
PR 16-AUG-2001; 2001US-312902P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;  
PI Vernet CAM, Li L, Shenoy S, Casman SJ;  
XX  
DR WPI; 2002-528447/56.  
DR N-PSDB; ABK97214.  
XX  
XX  
PT New G-protein coupled receptor polypeptides for treating or preventing  
PT cardiomyopathy, atherosclerosis, diabetes, multiple sclerosis, acquired  
PT immunodeficiency syndrome or cancer in humans  
XX  
PS Claim 1; Page 86; 110pp; English.  
XX  
CC The present invention relates to a new G-protein coupled receptor (GPCRX)  
CC polypeptide. The GPCR polypeptide, GPCR nucleic acid and antibody are  
CC useful for treating, preventing or alleviating a GPCR-associated  
CC disorder or a pathological state in a subject, particularly a human. In  
CC particular, the disorder is cardiomyopathy, atherosclerosis, diabetes, or  
CC a disorder related to cell signal processing and metabolic pathway  
CC modulation. The GPCR polypeptide and nucleic acid are also useful for  
CC diagnosing the presence of or predisposition to a disease associated  
CC with altered levels of GPCR, particularly cancer. The GPCR nucleic acid  
CC and polypeptide are especially useful in therapeutic or prophylactic

CC applications for disorders of the neuro-olfactory system, e.g. those  
CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding  
CC the protein is useful in gene therapy for treating the above conditions.  
CC Furthermore, the nucleic acids and polypeptides are useful in treating  
CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune  
CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's  
CC disease, multiple sclerosis or Albright hereditary osteodystrophy. These  
CC are also useful in developing powerful assay system for functional  
CC analysis of various human disorders, as well as in diagnostic  
CC applications. The present amino acid sequence represents one of a  
CC collection (AAU68134-AAU68153) of human GPCR proteins of the invention.  
XX  
SQ Sequence 315 AA;  
Query Match 99.18; Score 1620; DB 23; Length 315;  
Best Local Similarity 99.48; Pred. No. 3.9e-177;  
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVMVFYVALCGNVLLIFLIYLDAGLHTPM 60  
Db 1 MGRWVNSYTDGFFLLGIFSHSQTDLVLFSAVMVFYVALCGNVLLIFLIYLDAGLHTPM 60  
QY 61 YFFLSQLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFFVSLVSGELLLGLMA 120  
Db 61 YFFLSQLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFFVSLVSGELLLGLMA 120  
QY 121 YDHYVAVSHPLHYPILMNQVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSDHFFC 180  
Db 121 YDRYVAVSHPLHYPILMNQVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSDHFFC 180  
QY 181 EVQALLKLACADTSLFDTLLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
Db 181 EVQALLKLACADTSLFDTLLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
QY 241 TCSSHLTAVTLEFYGAAMFYLPRRYRAPSHDKVASIFYVTLTPLMPLIYSLRNGEVMG 300  
Db 241 TCSSHLTAVTLEFYGAAMFYLPRRYRAPSHDKVASIFYVTLTPLMPLIYSLRNGEVMG 300  
QY 301 ALRKGLDRCRIGSQH 315  
Db 301 ALRKGLDRCRIGSQH 315  
RESULT 4  
AAG71699  
ID AAG71699 standard; Protein; 315 AA.  
XX  
AC AAG71699;  
XX  
DT 30-JUL-2001 (first entry)  
XX  
DE Human olfactory receptor polypeptide, SEQ ID NO: 1380.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation.  
XX  
OS Homo sapiens.  
XX  
PN WO200127158-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 06-OCT-2000; 2000WO-US27582.  
XX  
PR 08-OCT-1999; 99US-0158615.  
PR 24-FEB-2000; 2000US-0184809.  
XX  
PA (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX

DR WPI; 2001-290713/30.  
XX  
XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX  
XX Claim 11; Page 860-861; 1857pp; English.  
XX  
CC The present sequence is an olfactory receptor which is encoded by  
CC one of a number of novel polynucleotides. The polynucleotides can be  
CC used in screening for olfactory agonists and antagonists. The methods  
CC allow for the determination of primary scents and the identification  
CC of the odour receptors used to detect these primary scents. The methods  
CC also enable determination of secondary scents and the identification of  
CC combinations of odour receptors that are involved in detecting such  
CC secondary scents. This enables the construction of a scent representation  
CC (also called a scent fingerprint or scent profile), which may be used to  
CC re-create and edit scents. Libraries of olfactory receptors are useful  
CC for determining the interaction pattern of a composition with the  
CC receptors, and can be used for determining differences in the olfactory  
CC faculties of different individuals.  
XX  
SQ Sequence 315 AA:  
  
Query Match 87.6%; Score 1432; DB 22; Length 315;  
Best Local Similarity 87.0%; Pred. No. 1.4e-155;  
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;  
  
OY 1 MGRWVNQSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLIYDAGLHTPM 60  
1 METWVNQSYTDGFFLLGIFSHSTADLVLFVSVMAVFTVALCGNVLLIFLIYMDPHLHTPM 60  
DB 1 YFFLSQLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGCIQIGLFFVSLVGSEGLLGMA 120  
61 YFFLSQLSLMDLMLVCTNVPKMAANFLSGRKSISFVCGCIQIGLFFVCLVGSEGLLGMA 120  
OY 121 YDHVAVSHPLHXPILMNQRCVLTGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
121 YDRYVAISHPLHXPILMNQRCVLTGSSWAFGIIDGVIQWVVMNFPYCGLRKVNHFEC 180  
DB 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALA 240  
181 EMLSLKLACVDTSLFEKVIACCVFMLLPFSIIVASYARILGTVLQMHSAQAMKKALA 240  
OY 241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
DB 241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
OY 301 ALRKGIDRCRIGSQH 315  
301 ALRKGIDRCRIGSQH 315  
DB 301 ALRKGIDRCRIGSQH 315

RESULT 5  
AAG71792  
ID AAG71792 standard; Protein: 315 AA.  
XX  
XX AAG71792;  
AC  
XX 30-JUL-2001 (first entry)  
DT  
XX  
DE Human olfactory receptor polypeptide, SEQ ID NO: 1473.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation.  
XX  
OS Homo sapiens.  
XX  
XX WO200127158-A2.  
PN  
XX 19-APR-2001.  
PD  
XX 06-OCT-2000; 2000WO-US27582.  
PF

XX  
XX 08-OCT-1999; 99US-0158615.  
PR 24-FEB-2000; 2000US-0184809.  
XX  
XX (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.  
PA  
XX  
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX  
XX WPI; 2001-290713/30.  
DR  
XX  
PT New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX  
XX Claim 11; Page 932; 1857pp; English.  
PS  
XX  
XX The present sequence is an olfactory receptor which is encoded by  
CC one of a number of novel polynucleotides. The polynucleotides can be  
CC used in screening for olfactory agonists and antagonists. The methods  
CC allow for the determination of primary scents and the identification  
CC of the odour receptors used to detect these primary scents. The methods  
CC also enable determination of secondary scents and the identification of  
CC combinations of odour receptors that are involved in detecting such  
CC secondary scents. This enables the construction of a scent representation  
CC (also called a scent fingerprint or scent profile), which may be used to  
CC re-create and edit scents. Libraries of olfactory receptors are useful  
CC for determining the interaction pattern of a composition with the  
CC receptors, and can be used for determining differences in the olfactory  
CC faculties of different individuals.  
XX  
SQ Sequence 315 AA:  
  
Query Match 87.6%; Score 1432; DB 22; Length 315;  
Best Local Similarity 87.0%; Pred. No. 1.4e-155;  
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;  
  
OY 1 MGRWVNQSYTDGFFLLGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLIYDAGLHTPM 60  
1 METWVNQSYTDGFFLLGIFSHSTADLVLFVSVMAVFTVALCGNVLLIFLIYMDPHLHTPM 60  
DB 1 YFFLSQLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGCIQIGLFFVCLVGSEGLLGMA 120  
61 YFFLSQLSLMDLMLVCTNVPKMAANFLSGRKSISFVCGCIQIGLFFVCLVGSEGLLGMA 120  
OY 121 YDHVAVSHPLHXPILMNQRCVLTGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
121 YDRYVAISHPLHXPILMNQRCVLTGSSWAFGIIDGVIQWVVMNFPYCGLRKVNHFEC 180  
DB 181 EVOALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALA 240  
181 EMLSLKLACVDTSLFEKVIACCVFMLLPFSIIVASYAHILGTVLQMHSAQAMKKALA 240  
OY 241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
DB 241 TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPLNPLIYSLRNGEVMG 300  
OY 301 ALRKGIDRCRIGSQH 315  
301 ALRKGIDRCRIGSQH 315  
DB 301 ALRKGIDRCRIGSQH 315

RESULT 6  
AAG72461  
ID AAG72461 standard; Protein: 315 AA.  
XX  
XX AAG72461;  
AC  
XX 30-JUL-2001 (first entry)  
DT  
XX  
DE Human OR-like polypeptide query sequence, SEQ ID NO: 2142.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW

KW scent profile; scent fingerprint; scent representation.  
XX OS Homo sapiens.  
XX PN WO200127158-A2.  
XX PD 19-APR-2001.  
XX PF 06-OCT-2000; 2000WO-US27582.  
XX PR 08-OCT-1999; 99US-0158615.  
XX PR 24-FEB-2000; 2000US-0184809.  
XX PA (DIGI-) DIGISCENTS.  
XX PA (YEDA ) YEDA RES & DEV CO LTD.  
XX PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX DR WPI; 2001-290713/30.  
XX PT New polynucleotides which encode polypeptides involved in olfactory  
XX PT sensation for identifying olfactory agonists and antagonists -  
XX PS Example 6; Page 1439-1440; 1857pp; English.  
XX CC The present sequence is a polypeptide encoded by one of 344 newly mined  
XX CC human genes. It was used as a query sequence in a database search of  
XX CC olfactory receptor (OR)-like sequences. The invention relates to isolated  
XX CC polynucleotides encoding polypeptides involved in olfactory sensation.  
XX CC The polynucleotides can be used in screening for olfactory agonists and  
XX CC antagonists. The methods allow for the determination of primary scents  
XX CC and the identification of the odour receptors used to detect these  
XX CC primary scents. The methods also enable determination of secondary scents  
XX CC and the identification of combinations of odour receptors that are  
XX CC involved in detecting such secondary scents. This enables the  
XX CC construction of a scent representation (also called a scent fingerprint  
XX CC or scent profile), which may be used to re-create and edit scents.  
XX CC Libraries of olfactory receptors are useful for determining the  
XX CC interaction pattern of a composition with the receptors, and can be  
XX CC used for determining differences in the olfactory faculties of different  
XX CC individuals.  
SQ Sequence 315 AA;  
Query Match 87.6%; Score 1432; DB 22; Length 315;  
Best Local Similarity 87.0%; Pred. No. 1.4e-155;  
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;  
QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVFSVAVMVFTVALCGNVLLIFLIYLDAGLHTPM 60  
Db 1 METWVNSYTDGFFLLGIFSHSTADLVFSVMAVFTVALCGNVLLIFLIYMDPHLHTPM 60  
QY 61 YFFLSQSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFVSLVSGELLLGLMA 120  
Db 61 YFFLSQSLMDLMLVCTNVPKMAANFLSGRKSISFVCGGIQIGFVCLVSGELLLGLMA 120  
QY 121 YDHYVAVSHPLHYPIILMNQVCLQITGSSNAFGIIDGVIQMVAAAGLPYCGSRSDHFFC 180  
Db 121 YDRYVAISHPLHYPIILMNQVCLQITGSSNAFGIIDGLIQMVVVMNPFYGLRKVNHFEC 180  
QY 181 EVQALLKLACADTSLEDTLLFACCVFEMLLPFSTIMASYACILGAVLRISAQAQWKALA 240  
Db 181 EMLSLKLACVDTSLEKVIYFACCVFEMLLPFSTIVASTARILGLTVLQMSAQAQWKALA 240  
QY 241 TCSSHLTAATLYFGAAMFVLRPRRYRADSHDKVASIFTVLTPLNPLIYSLRNGEVMG 300  
Db 241 TCSSHLTAATLYFGAAMFVLRPRRYRADSHDKVASIFTVLTPLNPLIYSLRNGEVMG 300  
QY 301 ALRKGIDRCRIGSQH 315  
Db 301 ALRKGIDRCRIGSQH 315

RESULT 7  
AAU95619  
ID AAU95619 standard; Protein; 315 AA.  
XX AC AAU95619;  
XX AC AAU95619;  
XX DT 02-JUL-2002 (first entry)  
XX DE Human olfactory and pheromone G protein-coupled receptor #106.  
XX KW Human; olfactory and pheromone G protein coupled; receptor;  
KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;  
KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
KW sterility; psychotic disorder; neurological disorder; anxiety;  
KW schizophrenia; manic depression; depression; axonal growth;  
KW menstrual cycle; appetite sexual motivation; sexual attraction;  
KW aggression.  
XX KW  
XX OS Homo sapiens.  
XX PN WO200224726-A2.  
XX PD 28-MAR-2002.  
XX PF 21-SEP-2001; 2001WO-BE00162.  
XX PR 22-SEP-2000; 2000EP-0870211.  
XX PA (CHEM-) CHEMCOM SA.  
XX PI Veitken A;  
XX DR WPI; 2002-330013/36.  
XX DR N-PSDB; ABK68506.  
XX PT Novel pheromone G-protein coupled receptor and receptor-derived  
XX PT agonists, antagonists or inhibitors useful in food or cosmetic products  
XX PT or in the treatment or prevention of neurological disorders such as  
XX PT anxiety and schizophrenia -  
XX PS Disclosure; Page 331-332; 833pp; English.  
XX CC The invention relates to olfactory and pheromone G-protein coupled  
XX CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
XX CC portion and its encoding polynucleotide. Also included are an agonist,  
XX CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector  
XX CC comprising the polynucleotide, a cell transformed by the vector, a non-  
XX CC human mammal comprising a partial or total deletion of the polynucleotide  
XX CC encoding the receptor and screening (detection and possibly, recovering)  
XX CC of compounds which are known or not known to be agonist, antagonists or  
XX CC inhibitors of natural compounds to the GPCR. The receptor-derived  
XX CC agonists, antagonists, inhibitors or compounds are used as an  
XX CC improvement, elimination or substitution of an existing taste and/or a  
XX CC fragrance of (or in) the food and/or cosmetic products. They can also be  
XX CC used in the preparation of medicament in the treatment and/or prevention  
XX CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
XX CC neurological disorders, including anxiety, schizophrenia, manic  
XX CC depression, depression, for promoting axonal growth, nerve cell  
XX CC connection and nerve regeneration for modulating male and female  
XX CC endocrine functions, hormone production and the menstrual cycle, for the  
XX CC prevention or the treatment by stimulation of several mammalian  
XX CC behaviours, such as stimulation or suppression of appetite, sexual  
XX CC motivation, sexual attraction, aggression and for promoting or  
XX CC suppressing chemical communication between organisms. The present  
XX CC sequence is a human olfactory and pheromone GPCR protein sequence.  
SQ Sequence 315 AA;  
Query Match 87.6%; Score 1432; DB 23; Length 315;  
Best Local Similarity 87.0%; Pred. No. 1.4e-155;  
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;  
QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVFSVAVMVFTVALCGNVLLIFLIYLDAGLHTPM 60

Db 1 METWVNSQSYTDGFFLLGIFSHSTADLVLFVSVMVFTVALCGNVLLIFLIYMDPHLHTPM 60  
QY 61 YFFLSQSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQIGFFVSLVSGEGLLGLMA 120  
Db 61 YFFLSQSLMDLMLVCTNVPKMAANFLSGRKSISFVGCIGIQIGLFCVLGSEGLLGLMA 120  
QY 121 YDHYVAVSHPLHYPIILMNORVCLQITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180  
Db 121 YDRYVAISHPLHYPIILMNORVCLQITGSSWAFGIIDGLIQWVVMNFPYCGLRKVNHFEC 180  
QY 181 EVQALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240  
Db 181 EMLSLKLACVDTSLFEKVIACCVFMLLPFSIIVASAHILGTVLQMHSAQAWKKALA 240  
QY 241 TCSSHLTAVTLFYGAMFMYLRRRYRAPSHDKVASIFVTYVLTPLMLNPLIYSLRNGEVMG 300  
Db 241 TCSSHLTAVTLFYGAMFMYLRRRYRAPSHDKVASIFVTYVLTPLMLNPLIYSLRNGEVMG 300  
QY 301 ALRKGDLRCRIGSOH 315  
Db 301 ALRKGDLRCRIGSOH 315

RESULT 8

AAE18656  
ID AAE18656 standard; Protein; 315 AA.

AAE18656;

17-MAY-2002 (first entry)

Human G-protein coupled receptor (GCREC-17).

Human: G-protein coupled receptor; GCREC-17; cell proliferative disorder; neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory; metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease; Pick's disease; Huntington's disease; Parkinson's disease; hypertension; atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic; osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS; anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening; transgenic animal; allergy; gene therapy; hepatotropic; immunosuppressant; nootropic; neuroprotective; cardiant; immunosuppressive; anorectic; virucide; receptor.

Homo sapiens.

Location/Qualifiers

1..24 /label= Signal\_peptide

25..315 /label= Human\_mature\_GCREC-17

26..46 /note= "Transmembrane domain"

98..119 /note= "Transmembrane domain"

201..220 /note= "Transmembrane domain"

WO200210387-A2.

07-FEB-2002.

25-JUL-2001; 2001WO-US23433.

27-JUL-2000; 2000US-221478P.  
03-AUG-2000; 2000US-223268P.  
21-AUG-2000; 2000US-227054P.  
08-SEP-2000; 2000US-231121P.  
13-SEP-2000; 2000US-232243P.  
15-SEP-2000; 2000US-232691P.  
22-SEP-2000; 2000US-235146P.

PA (INCY-) INCYTE GENOMICS INC.  
XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK;  
PI Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L;  
PI Graul RC, Warren BA, Lee EA, Ding L;  
XX WPI: 2002-188744/24.  
DR N-PSDB; AAD29683.

PT New human G-protein coupled receptor polypeptide for diagnosis,  
PT prevention and treatment of cell proliferative, neurological,  
PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and  
PT metabolic disorders

PS Claim 1; Page 135-136; 150pp; English.

XX The invention relates to novel human G-protein coupled receptors  
CC (GCREC) and their encoding polynucleotides. GCREC is useful as an  
CC immunogen for preparing monoclonal and polyclonal antibodies. GCREC is  
CC useful for diagnosing, treating and preventing a cell proliferative  
CC disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder  
CC (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's  
CC disease, Parkinson's disease), a cardiovascular disorder (e.g.,  
CC atherosclerosis, hypertension, myocardial infarction), a gastrointestinal  
CC disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/  
CC inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS),  
CC allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder  
CC (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is  
CC useful in a number of drug screening techniques, and to analyse the  
CC proteome of a tissue or cell type. GCREC is useful for creating knockin  
CC humanised animals or transgenic animals to model human diseases, in  
CC tissue or cell type, for detecting differences in the chromosomal  
CC location due to translocation, inversion, etc., among normal, carrier  
CC or affected individuals, and as hybridization probes for mapping  
CC naturally occurring genomic sequences. GCREC is useful in Southern or  
CC northern analysis, dot blot or other membrane-based technologies, in PCR  
CC technologies, in dipstick, pin, multiformat enzyme linked immunosorbant  
CC (ELISA)-like assays, and in microarrays utilising fluids or tissues from  
CC patients to detect altered GCREC expression. The present sequence is  
CC human GCREC-17.

XX Sequence 315 AA;

XX Query Match 87.6%; Score 1432; DB 23; Length 315;

XX Best Local Similarity 87.0%; Pred. No. 1.4e-155;

XX Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;

QY 1 MGRWVNSQSYTDGFFLLGIFSHQTDLVLFSAVWVFTVALCGNVLLIFLIYDAGLHTPM 60

Db 1 METWVNSQSYTDGFFLLGIFSHSTADLVLFVSVMVFTVALCGNVLLIFLIYMDPHLHTPM 60

QY 61 YFFLSQSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQIGFFVSLVSGEGLLGLMA 120

Db 61 YFFLSQSLMDLMLVCTNVPKMAANFLSGRKSISFVGCIGIQIGLFCVLGSEGLLGLMA 120

QY 121 YDHYVAVSHPLHYPIILMNORVCLQITGSSWAFGIIDGVIQWVAAMGLPYCGSRSDHFFC 180

Db 121 YDRYVAISHPLHYPIILMNORVCLQITGSSWAFGIIDGLIQWVVMNFPYCGLRKVNHFEC 180

QY 181 EVQALLKLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAWKKALA 240

Db 181 EMLSLKLACVDTSLFEKVIACCVFMLLPFSIIVASAHILGTVLQMHSAQAWKKALA 240

QY 241 TCSSHLTAVTLFYGAMFMYLRRRYRAPSHDKVASIFVTYVLTPLMLNPLIYSLRNGEVMG 300

Db 241 TCSSHLTAVTLFYGAMFMYLRRRYRAPSHDKVASIFVTYVLTPLMLNPLIYSLRNGEVMG 300

QY 301 ALRKGDLRCRIGSOH 315

Db 301 ALRKGDLRCRIGSOH 315

RESULT	9
AAU85223	
ID	AAU85223 standard; Protein; 315 AA.
XX	
AC	AAU85223;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	G-coupled olfactory receptor #84.
XX	
KW	Human; olfactory G-coupled receptor; sensory perception of odourant;
XX	
KM	odour composition; taste composition.
OS	Homo sapiens.
XX	
PN	WO200198526-A2.
XX	
PD	27-DEC-2001.
XX	
PF	22-JUN-2001; 2001WO-US20122.
XX	
PR	22-JUN-2000; 2000US-213812P.
XX	
PR	13-MAR-2001; 2001US-0804291.
XX	
PA	(SENO-) SENOMYX INC.
XX	
PI	Zozulya S, Stryer L;
XX	
DR	WPI; 2002-083330/11.
DR	N-PSDB; ABK37582.
PT	
PT	Representing sensory perception of one or more odourants for the
PT	identification and design of tastes and odours comprises providing a
PT	representative group of n olfactory receptors -
XX	
PS	Claim 1; Page 96; 182pp; English.
XX	
CC	The invention relates to a method of representing sensory perception of
CC	one or more odourants. The method comprises: (a) providing a
CC	representative class of n olfactory receptors or ligand binding domains
CC	(LBD) of these receptors; (b) measuring values X1 to Xn representative
CC	of at least one activity of one or more odourants selected from:
CC	(i) binding one or more odourants to the LBD of at least one of the n
CC	olfactory receptors; (ii) activating at least one of the n
CC	olfactory receptors with the one or more odourants; and (iii) blocking at
CC	least one of the n olfactory receptors with the one or more odourants;
CC	and (c) generating a representation of sensory perception from the values
CC	X1 to Xn. The representation of the sensory perception of odourants is
CC	useful for the design and formulation of odour and taste compositions.
CC	AAU85140-AAU85393 represent human olfactory G-coupled receptor amino acid
CC	sequences of the invention.
XX	
SQ	Sequence 315 AA;
Query Match	87.6%; Score 1432; DB 23; Length 315;
Best Local Similarity	87.0%; Pred. No. 1.4e-155;
Matches 274; Conservative	13; Mismatches 28; Indels 0; Gaps 0;
DQ	
1	MGRWNQSYTDGFFLLGFISHSQDPDLVFSAVMVVFVALCGNVLLIFLIYLDAGLHTPM 60
1	MEFWNQSYTDGFFLLGFISHSSTADLVLFVSVMAMFVALCGNVLLIFLIYMDPHLTPM 60
DQ	
61	VFELSQLSLMDIMLYCNIVPKMAANFLSGRKSISFVGCGIQIGFEVSLVGEGLLGMA 120
61	VFFLSQLSLMDIMLYCTNVPKMANFLSGRKSISFVGCGIQIGLFVCLVGSEGLLGMA 120
DQ	
121	YDHYVAVSHPLAHPILMNORVCLOITGSSWAFGIIDGVIQMVAAGLPYCGRSVDHFFC 180
	:
121	YDRYAIAISHPLAHPILMNQRVCLOITGSSSWAFGIIDGLIQMVVMNPYCGLRKVNHFFC 180
DQ	
181	EVOALLKLACADTSLEFDTLTFACCVFMLLPFSIIASYSACILGAVLRIRSAQAWKKALA 240
	: :                       : :

Db	181	EMSLKLACVDTSLFEEKVIFACCVFMLLFPSPSIIVASYAHILGIVLQMHSAQAWKKALA	240
QY	241	TCSSHLTAATVLFYGAAMFMYLRPRRYRAPSHDKVASIFYYTLTPMLNPLIYSLRNGEVMG	300
Db	241	TCSSHLTAATVLFYGAAMFMYLRPRRYRAPSHDKVASIFYYTLTPMLNPLIYSLRNGEVMG	300
QY	301	ALRKGLDRCRIGSQH	315
Db	301	ALRKGLDRCRIGSQH	315
RESULT 10			
ID	AAE16175	standard; Protein; 271 AA.	
XX	AAE16175;		
AC			
XX	26-MAR-2002	(first entry)	
DT			
XX			
DE		Human G-protein coupled receptor 6 (GCREC-6) protein.	
XX			
KW		Human; G-protein coupled receptor 6; cell proliferative disorder;	
KW		atheriosclerosis; hepatitis; cancer; neurological disorder; epilepsy;	
KW		Alzheimer's disease; Parkinson's disease; cardiovascular disorder;	
KW		atherosclerosis; hypertension; myocardial infarction; peptic ulcer;	
KW		gastrointestinal disorder; dysphagia; anorexia; autoimmune disorder;	
KW		acquired immune deficiency syndrome; inflammatory disorder; infection;	
KW		Addison's disease; allergy; Grave's disease; metabolic disorder; AIDS;	
KW		diabetes; obesity; osteoporosis; gene therapy; GCREC-6.	
XX			
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	Domain	26..46	
FT		/note= "Transmembrane domain"	
FT	Domain	42..155	
FT		/note= "7 transmembrane receptor domain"	
FT	Domain	111..127	
FT		/note= "G-protein coupled receptor motif"	
FT	Domain	223..247	
FT		/note= "7 transmembrane receptor domain"	
XX			
PN	WO200187937-A2.		
XX			
PD	22-NOV-2001.		
XX			
PF	17-MAY-2001; 2001WO-US16285.		
XX			
PR	18-MAY-2000; 2000US-205628P.		
PR	22-MAY-2000; 2000US-206222P.		
PR	25-MAY-2000; 2000US-207566P.		
PR	02-JUN-2000; 2000US-208834P.		
PR	02-JUN-2000; 2000US-208861P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Patterson C, Lu DAM, Thornton M, Lu Y, Tribouley CM, Graul R;		
PI	Khan FA, Gandhi AR, Wadia NK, Nguyen DB, Yue H, Hafalia A;		
PI	Elliott VS, Lal P, Reddy R, Kallick DA, Tang TY, Au-Young J;		
XX			
DR	WPI; 2002-089844/12.		
DR	N-PSDB; AAD26374.		
XX			
PT	Novel G-protein coupled receptors and polynucleotides useful for		
PT	diagnosis, treatment and prevention of disorders of cell proliferation,		
PT	neurological, cardiovascular, metabolic disorders and viral infections		
XX			
PS	Claim 1; Page 108; 115pp; English.		
XX			
CC	The invention relates to human G-protein coupled receptor (GCREC)		
CC	polypeptides and polynucleotides. GCREC polypeptides are useful for		
CC	screening compounds that modulate their activity. They are useful in		

CC the diagnosis, prevention and treatment of disorders which include  
CC cell proliferative disorders such as arteriosclerosis, hepatitis,  
CC myelofibrosis, psoriasis and cancer including adenocarcinoma, leukaemia,  
CC lymphoma; neurological disorders such as epilepsy, ischaemic  
CC cerebrovascular disease, Alzheimer's disease, Pick's disease, dementia,  
CC parkinson's disease, ataxias, multiple sclerosis, bacterial and viral  
CC meningitis, Creutzfeldt-Jakob disease, schizophrenic disorders, amnesia;  
CC cardiovascular disorders such as arteriovenous fistula, atherosclerosis,  
CC hypertension, vascular tumours, myocardial infarction, hypertensive  
CC heart disease, infective endocarditis, cardiomyopathy, myocarditis;  
CC gastrointestinal disorders such as dysphagia, peptic oesophagitis,  
CC emesis, anorexia, nausea, peptic ulcer, cholelithiasis, diarrhoea,  
CC constipation, acquired immune deficiency syndrome (AIDS), hepatic  
CC encephalopathy; autoimmune/inflammatory disorders such as Addison's  
CC disease, allergies, spondylitis, amyloidosis, anaemia, asthma, contact  
CC dermatitis, Crohn's disease, diabetes mellitus, Goodpasture's syndrome,  
CC emphysema, Grave's disease, gout, multiple sclerosis, rheumatoid  
CC arthritis, systemic lupus erythematosus, uveitis, viral, bacterial,  
CC fungal, parasitic, protozoal and helminthic infections and trauma;  
CC metabolic disorders such as diabetes, obesity and osteoporosis; and  
CC viral infections such as infection caused by viral agent classified as  
CC adenovirus, arenavirus, bunyavirus. Polynucleotides of the invention  
CC are useful as probes for assessing toxicity of test compounds. They are  
CC also used in gene therapy. The present sequence is human G-protein  
CC coupled receptor 6 (GCRC-6) protein.

XX SQ Sequence 271 AA;

Query Match 83.2%; Score 1360; DB 23; Length 271;  
Best Local Similarity 85.4%; Pred. No. 2.2e-147;  
Matches 269; Conservative 0; Mismatches 2; Indels 44; Gaps 1;

QY 1 MGRWNQSYTDGFFLLGIFSHSQTDLVLSAVMVVETVALCGNVLLIFLLYLDAGLHTPM 60  
Db 1 MGRWNQSYTDGFFLLGIFSHSQTDLVLSAVMVVETVALCGNVLLIFLLYLDAGLHTPM 60  
QY 61 YFELSLSMDLMLVNCNIVPKMANFLSGRKSISFVGGCIGIQIGFFVSLVSGELLLGLMA 120  
Db 61 YFELSLSMDLMLVNCNIVPKMANFLSGRKSISFVGGCIGIQIGFFVSLVSGELLLGLMA 120  
QY 121 YDHYVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSVDHFFC 180  
Db 121 YDHYVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSVDHFFC 180  
QY 121 YDRYVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSVDHFF- 179  
Db 121 YDRYVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSVDHFF- 179  
QY 181 EVQALLKACADTSLFTLLFACCVFMLLPISIMASYACILGAVLRIRSAQAWKKALA 240  
Db 181 EVQALLKACADTSLFTLLFACCVFMLLPISIMASYACILGAVLRIRSAQAWKKALA 240  
QY 180 -----WAVLRIRSAQAWKKALA 196  
Db 180 -----WAVLRIRSAQAWKKALA 196  
QY 241 TCSSHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMG 300  
Db 197 TCSSHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMG 256  
QY 301 ALRKGLDRCRIGSQH 315  
Db 257 ALRKGLDRCRIGSQH 271

RESULT 11  
AAU24603  
ID AAU24603 standard; Protein; 252 AA.  
XX AC AAU24603;  
XX DT 18-DEC-2001 (first entry)  
XX DE Human olfactory receptor AOLFR94.  
XX KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;  
XX KM food additive; cosmetic; fragrance; pharmaceutical additive.  
XX OS Homo sapiens.  
XX PN WO200168805-A2.

XX 20-SEP-2001.  
XX PF 13-MAR-2001; 2001WO-US07771.  
XX 13-MAR-2000; 2000US-0188914.  
PR 24-MAR-2000; 2000US-0192033.  
PR 12-APR-2000; 2000US-0198474.  
PR 24-APR-2000; 2000US-0199335.  
PR 26-MAY-2000; 2000US-0207702.  
PR 23-JUN-2000; 2000US-0213849.  
PR 16-AUG-2000; 2000US-0226534.  
PR 07-SEP-2000; 2000US-0230732.  
PR 07-FEB-2001; 2001US-0266862.  
XX PA (SENO-) SENOMYX INC.  
XX PI Zozulya S;  
XX DR WPI; 2001-570867/64.  
DR N-PSDB; AAS42296.

Nucleic acids encoding human olfactory G protein-coupled receptors,  
useful for screening for compounds involved in olfactory sensation,  
where the compounds can be used in the food, pharmaceutical and  
cosmetic industries to customise odours

Claim 60; Page 123; 319pp; English.

The invention relates to nucleic acids encoding human olfactory  
receptors, OR, (a G protein-coupled receptor, GPCR). The OR's  
specifically recognise molecules, odourants, that elicit specific  
olfactory sensation. The human olfactory receptors and polynucleotides  
encoding them are useful for screening a library of chemical compounds  
for compounds that are involved in olfactory sensation. Modulators of  
their activity are useful for pharmacological and genetic modulation of  
olfactory signalling pathways. Therefore, they can be used in the food,  
pharmaceutical and cosmetic industries to customise odours and  
fragrances. The present sequence is a human olfactory receptor of the  
invention.

XX SQ Sequence 252 AA;

Query Match 70.4%; Score 1150; DB 22; Length 252;  
Best Local Similarity 87.3%; Pred. No. 2.4e-123;  
Matches 220; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 64 LSQSLMDLMLVNCNIVPKMANFLSGRKSISFVGGCIGIQIGFFVSLVSGELLLGLMAYDH 123  
Db 1 LSQSLMDLMLVNCNIVPKMANFLSGRKSISFVGGCIGIQIGFFVCLVSGELLLGLMAYDR 60  
QY 124 YVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGVIQMVAAAGLPYCGSRSVDHFFCEVQ 183  
Db 61 YVAVSHPLHYPILMNQRVCLQITGSSWAFGIIDGLIQMVVVMNFPYCGLRKVNHFCEML 120  
QY 184 ALLKLCACADTSLFTLLFACCVFMLLPFSITMASYACILGAVLRIRSAQAWKKALATCS 243  
Db 121 SLKLCACVDTSLEKVIIFACCVFMLLPFSITVASYAHILGTVLQMHSAQAWKKALATCS 180  
QY 244 SHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMGALR 303  
Db 181 SHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMGALR 240  
QY 304 KGLDRCRIGSQH 315  
Db 241 KGLDRCRIGSQH 252

RESULT 12  
ABP61132  
ID ABP61132 standard; Protein; 263 AA.  
XX AC ABP61132;



XX 18-SEP-2002 (first entry)  
 DT XX  
 DE Human GPCR protein #4.  
 XX  
 KW Human; G-protein coupled receptor; GPCR; cardiac; antidiabetic; antidiabetic; hypotensive; anorectic; cytosolic; immunosuppressive; neuroprotective; nootropic; antiparkinsonian; antilipemic; gene therapy; cardiomyopathy; atherosclerosis; cell signal processing; diabetes; metabolic pathway; hypertension; obesity; cancer; Alzheimer's disease; graft versus host disease; Parkinson's disorder; dyslipidaemia.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200255557-A2.  
 PN  
 XX 18-JUL-2002.  
 PD  
 XX 18-DEC-2001; 2001WO-US49111.  
 PF  
 XX 18-DEC-2000; 2000US-256635P.  
 PR 21-DEC-2000; 2000US-257876P.  
 PR 04-JAN-2001; 2001US-259743P.  
 PR 10-JAN-2001; 2001US-260718P.  
 PR 12-JAN-2001; 2001US-261498P.  
 PR 24-JAN-2001; 2001US-263689P.  
 PR 08-FEB-2001; 2001US-267464P.  
 PR 22-FEB-2001; 2001US-271021P.  
 PR 14-MAR-2001; 2001US-275946P.  
 PR 23-MAR-2001; 2001US-278150P.  
 PR 16-AUG-2001; 2001US-312902P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;  
 PI Vernet CAM, Li L, Shenoy S, Casman SJ, Gusev V;  
 XX  
 DR WPI; 2002-500839/53.  
 DR N-PSDB; ABQ88046.  
 XX  
 PT New G-protein coupled receptor polypeptides, encoding nucleic acids and  
 PT immunospecific antibodies, useful for treating cardiomyopathy,  
 PT atherosclerosis, diabetes or disorder related to cell signal processing  
 PT and metabolic pathway modulation -  
 PT  
 XX  
 PS Claim 1; Page 86; 146pp; English.  
 XX  
 CC The invention relates to a novel nucleic acid molecule comprising a  
 CC nucleic acid sequence encoding a G-protein coupled receptor (GPCR)  
 CC polypeptide. The proteins of the invention have cardiac,  
 CC antidiabetic, antidiabetic, hypotensive, anorectic, cytosolic,  
 CC immunosuppressive, neuroprotective, nootropic, antiparkinsonian, and  
 CC antilipemic activity. The nucleic acids of the invention may have a use  
 CC in gene therapy. The sequences of the invention are useful for treating  
 CC or preventing a GPCR-associated disorder, preferably cardiomyopathy,  
 CC atherosclerosis or a disorder related to cell signal processing and  
 CC metabolic pathway modulation. An antibody of the invention is useful for  
 CC treating diabetes or a disorder related to cell signal processing and  
 CC metabolic pathway modulation. GPCR-associated disorders include  
 CC hypertension, obesity, cancer, graft versus host disease, Alzheimer's  
 CC disease, Parkinson's disorder and dyslipidaemia. The sequences shown in  
 CC ABQ88043-ABQ88093 encode the GPCR proteins of the invention, shown in  
 CC ABP61129-ABP61179.  
 CC  
 XX  
 SQ Sequence 263 AA;

Query Match 70.1%; Score 1146; DB 23; Length 263;  
 Best Local Similarity 99.5%; Pred. No. 7.3e-123;  
 Matches 221; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGRVNSQITDGFLLGIFSHSQTDLVLSAVMVVFTVALCGNVLLIFLIYLDAGLHTPM 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 MGRVNSQITDGFLLGIFSHSQTDLVLSAVMVVFTVALCGNVLLIFLIYLDAGLHTPM 60

QY 61 YFFLSQSLMDLMLVCNIVPKMANFLSGRKSISFVCGIQIGFFVSLVSGEGLLGLMA 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 61 YFFLSQSLMDLMLVCNIVPKMANFLSGRKSISFVCGIQIGFFVSLVSGEGLLGLMA 120

QY 121 YDHYVAVSHPLHPILMNQVCLOITGSSWAFCIIDGVIQMYAAMGLPYCGSRVDHFFC 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 121 YDRYVAVSHPLHPILMNQVCLOITGSSWAFCIIDGVIQMYAAMGLPYCGSRVDHFFC 180

QY 181 EVQALLKLACADTSLFDTLFRACCVFMLLPSTIIMASYACI 222  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 181 EVQALLKLACADTSLFDTLFRACCVFMLLPSTIIMASYACI 222

RESULT 13  
 AAY96666  
 ID AAY96666 standard; Protein; 223 AA.  
 XX  
 AC AAY96666;  
 XX  
 DT 26-SEP-2000 (first entry)  
 XX  
 DE Murine olfactory receptor ligand-binding region 5.  
 XX  
 KW Degenerate primer; ligand-binding region; olfactory receptor;  
 KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;  
 KW human rhodopsin receptor; odorant; toxicity; olfactory response.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200035274-A1.  
 XX  
 PD 22-JUN-2000.  
 XX  
 PF 17-DEC-1999; 99WO-US30221.  
 XX  
 PR 17-DEC-1998; 98US-0112605.  
 XX  
 PA (UJJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 XX  
 PI Reed RR, Krautwurst D, Yau KW;  
 XX  
 DR WPI; 2000-431471/37.  
 DR N-PSDB; AAA51179.  
 XX  
 PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.  
 PT for producing receptor libraries used in e.g. screening odorants for  
 PT toxicity  
 PT  
 XX  
 PS Claim 18; Page 32; 61pp; English.  
 XX  
 CC AAY96662-80 are murine odorant/ligand binding regions of olfactory  
 CC receptors. The coding sequences were amplified using degenerate primers  
 CC shown in AAA51163-64.  
 CC The primers can be used to generate a library of OR-LBR comprising the  
 CC transmembrane domains (TMD) II-VI, III-VI, IV-VII or III-VII. Chimeric  
 CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an  
 CC N-terminal plasma membrane translocation domain (especially the sequence  
 CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain  
 CC was initially derived from the N-terminus of the human rhodopsin  
 CC receptor. LBR produced by amplification with the primers are useful for  
 CC generating new odorants, to screen for toxicity or therapeutic activity  
 CC in odorants, and altering an animal's olfactory response (claimed).  
 CC  
 XX  
 SQ Sequence 223 AA;

Query Match 63.9%; Score 1044; DB 21; Length 223;  
 Best Local Similarity 89.7%; Pred. No. 3e-111;  
 Matches 200; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 65 SOLSLMDLMLVCNIVPKMANFLSGRKSISFVCGIQIGFFVSLVSGEGLLGLMAYDHY 124  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 SOLSLMDLMLVCNIVPKMANFLSGRKSISFVCGIQIGFFVSLVSGEGLLGLMAYDHY 60

```
OY 125 VAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIOMVAAMGLPYCGSRSDVDFCEVQA 184
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 VAISHPLHYPIILMSQKYCLOIAGSSWAFGILDGIOMVAAMSLPYCGSRVIDHFCEVPA 120

OY 185 LKTLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALATCSS 244
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LKTLACADTSLFDTLFACCVFMLLPFSIIVTSYARILGAVLRMHSAQSRKKALATCSS 180

OY 245 HLTAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYTVLTPMLN 287
||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 HLTAVSLFYGAAMFYLRPRRYRAPSHDKVVSIFYTVLTPMLN 223
```

## RESULT 14

```
AAV96670
ID AAV96670 standard; Protein; 223 AA.
XX
AC AAV96670;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region 9.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response.
XX
OS Mus musculus.
XX
PN WO200035274-A1.
XX
PD 22-JUN-2000.
XX
PE 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
DR N-PSDB; AAA51183.
XX
PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
PT toxicity
XX
PS Claim 18; Page 33; 61pp; English.
XX
XX AAV96662-80 are murine odorant/ligand binding regions of olfactory
CC receptors. The coding sequences were amplified using degenerate primers
CC shown in AAA51163-64.
CC The primers can be used to generate a library of OR-LBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AAV96658), a first TMD and the OR-LBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. LBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
```

SQ Sequence 223 AA;

Query Match 63.9%; Score 1044; DB 21; Length 223;  
Best Local Similarity 89.7%; Pred. No. 3e-111;  
Matches 200; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```
OY 65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIOIGFVSLVSGELLLGLMAYDHY 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 SOLSLMDLMLVCNIVPKMAVNFLSGRKSISFAGCGIOIGFVSLVSGELLLGLMAYDRY 60
```

```
OY 125 VAVSHPLHYPIILMNQRYCLOITGSSWAFGIIDGVIOMVAAMGLPYCGSRSDVDFCEVQA 184
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 VAISHPLHYPIILMSQKYCLOIAGSSWAFGILDGIOMVAAMSLPYCGSRVIDHFCEVPA 120

OY 185 LKTLACADTSLFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKKALATCSS 244
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LKTLACADTSLFDTLFACCVFMLLPFSIIVTSYARILGAVLRMHSAQSRKKALATCSS 180

OY 245 HLTAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYTVLTPMLN 287
||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 HLTAVSLFYGAAMFYLRPRRYRAPSHDKVVSIFYTVLTPMLN 223
```

## RESULT 15

```
AAV96668
ID AAV96668 standard; Protein; 223 AA.
XX
AC AAV96668;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region 7.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response.
XX
OS Mus musculus.
XX
PN WO200035274-A1.
XX
PD 22-JUN-2000.
XX
PE 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
DR N-PSDB; AAA51181.
XX
PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
PT toxicity
XX
PS Claim 18; Page 32-33; 61pp; English.
XX
XX AAV96662-80 are murine odorant/ligand binding regions of olfactory
CC receptors. The coding sequences were amplified using degenerate primers
CC shown in AAA51163-64.
CC The primers can be used to generate a library of OR-LBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AAV96658), a first TMD and the OR-LBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. LBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
```

SQ Sequence 223 AA;

Query Match 63.6%; Score 1040; DB 21; Length 223;  
Best Local Similarity 89.2%; Pred. No. 8.5e-111;  
Matches 199; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```
OY 65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIOIGFVSLVSGELLLGLMAYDHY 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 SOLSLMDLMLVCNIVPKMAVNFLSGRKSISFAGCGIOIGFVSLVSGELLLGLMAYDRY 60
```

QY	125	VASHPLHLPILMNQVCLQTIGSSNAFGIIDDVTQMVAAMGLPYCGSRSDHFFCEVQA	184
		:     :     :     :     :     :     :	
Db	61	VAISHPLHPILMSQKVCLOIAGSSWAFGLDGIITQMVAAMSLPYCGSRYIDHFCEVPA	120
QY	185	LKLACADTSLEDTLLFACCVFMLLPFSIIIMASYACILGAVLRIRSAQAMKALATCSS	244
		:     :     :     :     :     :     :	
Db	121	LKLACADTSLEDTLLFACCVFMLLPFSIIIVTSARILGTVLRMHSAQSRRKALATCSS	180
QY	245	HLTAVTLFYGAAMFMLRPRRYRAPSHDKVASIFVTVLTPLN	287
		:     :     :     :     :     :     :	
Db	181	HLTAVSLFYGAAMFIYLRRRYRAPSHDKVASIFTVLTPLN	223

Search completed: April 28, 2003, 14:00:26  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:59:16 ; Search time 18 Seconds  
(without alignments)  
1682.352 Million:cell updates/sec

Title: US-09-634-109D-2  
Perfect score: 1634  
Sequence: 1 MGRWVNQSYDGFLLGIFS.....GEVWGLRKGLDRCRIGSQH 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	44.2	316	2 A57069	olfactory receptor
2	715	43.8	315	2 JC5836	olfactory receptor
3	696	42.6	312	2 A46247	olfactory receptor
4	654.5	40.1	312	2 A48413	probable olfactory
5	631	38.6	305	2 S29711	olfactory factor O
6	629	38.5	327	2 F23701	olfactory receptor
7	608	37.2	309	1 S51356	olfactory receptor
8	607	37.1	318	2 JC5202	chemoreceptor TB64
9	601	36.8	320	2 S20573	olfactory receptor
10	595	36.4	315	2 JC5201	chemoreceptor TB56
11	591.5	36.2	313	2 S20571	olfactory receptor
12	591	36.2	313	2 B23701	olfactory receptor
13	588	36.0	311	2 JC5200	chemoreceptor TB33
14	583	35.7	314	2 S29707	olfactory receptor
15	578	35.4	319	2 JC5624	olfactory receptor
16	573	35.1	304	2 S29709	olfactory receptor
17	571	34.9	333	2 A23701	olfactory receptor
18	570	34.9	314	2 S20572	olfactory receptor
19	570	34.9	314	2 A37286	olfactory receptor
20	568	34.8	314	2 H23701	olfactory receptor
21	562	34.4	307	2 S29710	olfactory receptor
22	560	34.3	312	2 S29708	olfactory receptor
23	558	34.1	310	2 E23701	olfactory receptor
24	552	33.8	315	2 JC4658	olfactory receptor
25	545.5	33.4	317	2 D23701	olfactory receptor
26	533	32.6	222	2 D40745	odorant receptor (
27	531	32.5	312	2 I23701	olfactory receptor
28	522	31.9	312	2 G23701	olfactory receptor
29	516	31.6	157	2 S58073	probable olfactory

30	501.5	30.7	312	2 A46750	olfactory receptor
31	490	30.0	311	2 C23701	olfactory receptor
32	483	29.6	157	2 S58037	probable olfactory
33	481	29.4	157	2 S58029	probable olfactory
34	479	29.3	222	2 B40745	odorant receptor (
35	471	28.8	264	2 PC4369	olfactory receptor
36	441	27.0	157	2 S58033	probable olfactory
37	437	26.7	157	2 S58019	probable olfactory
38	424	25.9	234	2 S29000	G protein-coupled
39	422	25.8	157	2 S58011	G protein-coupled
40	413	25.3	234	2 S28998	olfactory receptor
41	411.5	25.2	215	2 I38473	olfactory receptor
42	404	24.7	216	2 I38480	olfactory receptor
43	397.5	24.3	232	2 S29001	G protein-coupled
44	396.5	24.3	328	2 G45774	odorant receptor 2
45	396	24.2	216	2 I38481	olfactory receptor

ALIGNMENTS

RESULT 1					
A57069					
olfactory receptor FAT11 - human					
C:Species: Homo sapiens (man)					
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 26-Aug-1999					
C:Accession: A57069					
R:Fan, W.; Liu, Y.C.; Parimoo, S.; Weissman, S.M.					
Genomics 27, 119-123, 1995					
A:Title: Olfactory receptor-like genes are located in the human major histocompatibil					
A:Reference number: A57069; MUID:95394447; PMID:7665158					
A:Accession: A57069					
A:Status: preliminary					
A:Molecule type: mRNA					
A:Residues: 1-316 <FAN>					
A:Cross-references: GB:L35475; NID:g1041044; PIDN:AAB36567.1; PID:g601919					
C:Genetics:					
A:Gene: GDB:FAT11; OLFER2					
A:Cross-references: GDB:1323249; OMIM:600578					
A:Map position: 6p21.3-6p21.3					
C:Superfamily: olfactory receptor OR14					
Query Match					
Best Local Similarity 44.2%; Score 723; DB 2; Length 316;					
Matches 142; Conservative 52; Mismatches 111; Indels 0; Gaps 0;					
QY 6 NOSYTDGFLLGIFSHQTDLVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPMYFFLS 65					
Db 3 NOSSTGFLLLGFESEHGLGRTLFEVDVITSYLLTVGNTLITLLSALDTKLSPMYFFLS 62					
QY 66 QLSLMDLMLVCNIVPKMANFLSGRKSTISFVGCGIQIGFEVSLVGSSEGLLGLMAYDHY 125					
Db 63 NLSFLDLCTTSCVPQMLANLWGPVKTTISFLDCSVQIFLFLSGTTECLMKVMAFDHY 122					
QY 126 AVSHPLHYPILMNQVCLQITGSSWAFGIIDGVIQVAAAGLPYCGSRSDHFFCEVOAL 185					
Db 123 AVCQPLHYATIIHPRLCQLASVAVIGLVSVQTPSTLHLFCPPDRQVDDFVCEVPAL 182					
QY 186 LKLACADTSLFDTLFACCVFMLLPESIIMASYACIIGAVLRIRSAQAWKKALATGSSH 245					
Db 183 IRLSCEDTSYNEIQVAVSVFILVWPLSLILVSYGALTWAVLRINSATMARKAFTGSSH 242					
QY 246 LTVAVTFYGAAMFMVLRPRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMGALRKG 305					
Db 243 LTVVTFYSSVIAVYIQPNPYAQGRGKEFGLFYAVGTPSLNPLVYTLRNKEIKRALRRL 302					
QY 306 LDRCR 310					
Db 303 LGKER 307					
RESULT 2					
JC5836					

olfactory receptor-like protein 2 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 24-Nov-1999  
 C:Accession: J05836  
 R:Blache, P.; Gros, L.; Salazar, G.; Bataille, D.  
 Biochem. Biophys. Res. Commun. 242, 669-672, 1998  
 A>Title: Cloning and tissue distribution of a new rat olfactory receptor-like (OL2).  
 A:Reference number: J05836; MUID:98125551; PMID:9464275  
 A:Accession: J05836  
 A:Status: nucleic acid sequence not shown  
 A:Status: molecule type: DNA  
 A:Molecule type: DNA  
 A:Residues: 1-315 <BLA>  
 A:Cross-references: GB:AF029357; MID:g2570934; PIDN:AAC39969.1; PID:g2570935  
 C:Comment: This protein is implicated in white cell muturation and migration.  
 C:Superfamily: olfactory receptor OR14  
 F:28-53/Domain: transmembrane #status predicted <TM1>  
 F:64-86/Domain: transmembrane #status predicted <TM2>  
 F:95-123/Domain: transmembrane #status predicted <TM3>  
 F:147-165/Domain: transmembrane #status predicted <TM4>  
 F:203-229/Domain: transmembrane #status predicted <TM5>  
 F:241-264/Domain: transmembrane #status predicted <TM6>  
 F:277-293/Domain: transmembrane #status predicted <TM7>

Query Match	43.8%;	Score 715;	DB 2;	Length 315;
Best local Similarity	47.5%;	Pred. No. 1.4e-53;		
Matches 143;	Conservative 54;	Mismatches 104;	Indels 0;	Gaps 0;

QY	6	NOSYTDGEFLLGIFSHSQTDVLVESAVMVFTVALCGNVLLIFLIYDAGLHTPMYFFLS	65
D8	8	NCSVSEFILLGFSSSEQIRMALFIEPFLLLYMVTLLGNGLIVALIYDSRLHTPMYFFLS	67
QY	66	QLSLMDMLVCNIVPKMAANFLSGRRSISFVGCGIQIGFEVSVLGSEGLLGLMAYDHY	125
D8	68	ILSLVMSYVTTTVPQMLVNVMCPKRITSWGACVAOMEIFLVGIACCVLYAIMADRYI	127
QY	126	AVSHPLHPILMNORYCLOITGSSWAFGITDGVICQVAAAMGLPYCGSRSDHFFCEVOAL	185
D8	128	AICEPLHYSVLMSRLOCAMKMTICSSISVTGALITYFTFMRLPYCGBPYNHFCEVPAV	187
QY	186	LKLACADTSLEFDTLFACCVEFMLLPFSIIMASYACILGAVLRIRSAQAAMKKALATCSSH	245
D8	188	LKLACADTSFNDRDLDFILGFVLLVPLSLILASACYFSILRIRSSGRLKSFTSCASH	247
QY	246	LTAVTLEYGAMFMYLPRRRYRAPSHDKVASIFYTVLTPLMLNPILIYSLNGEVMGALKRG	305
D8	248	ITVVTMEYGPAWMVMYMRGSWYDPERDKLALLFYNNVSAFLNPITIYSLRNKDVKAEMKY	307
QY	306	L 306	
D8	308	L 308	

### RESULT 3

olfactory receptor OR3 - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-Aug-1999  
 C/Accession: A46247  
 R/Nef, P.; Hermans-Borgmeyer, I.; Artieres-Pin, H.; Beasley, L.; Dionne, V.E.; Heinemann  
 Proc. Natl. Acad. Sci. U.S.A. 89, 8948-8952, 1992  
 A/Title: Spatial pattern of receptor expression in the olfactory epithelium.  
 A/Reference number: A46247; MUID:93028384; PMID:1384038  
 A/Accession: A46247  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-312 <NEF>  
 A/Cross-references: GB:M84005; NID:g200153; PIDN:AAA39862.1; PID:g200154  
 A/Note: sequence extracted from NCBI backbone (NCBIP:115362)  
 C/Superfamily: olfactory receptor OR14

Query Match	42.68;	Score 696;	DB 2;	Length 312;
Best Local Similarity	45.28;	Pred. No. 5.9e-52;		
Matches 138; Conservative	56;	Mismatches 111;	Indels 0;	Gaps 0;

```

QY      6 NOSTDGEFFLLGIFSHSQTDLVLESVMMVFETVALCGNVLIFLLYLDAGLHTPMYFFLS 65
      : | : | : | : : : : : : | | : | | | | | | | | |
Db      5 SNSSSGTFILMGVSDPHLEIIEFAVILASLYLLTLVGNLTIILSLDLARLHTPMYFFLS 64

QY      66 QLSLMDLMLVCNIVPKMAANFLSGRKSSISFVGCGIQIGFEFVSLVGSEGLLGMAYDHYV 125
      | | : | : | : | : | : | : | : | : | : | : | : |
Db      65 NLSSDLAFTTSVPOMLKNLMBDKTISYGGCVTQLYFVLWGATECILLVMAFDRYV 124

QY      126 AVSHPLHYPILMNORVCLQITGSSWAFGIIDGVIQMVAAAMGLPYCGSRSDHFECEVOAL 185
      | | | | : | : | : | : | : | : | : | : | : | : | :
Db      125 AVCRLHYMTVMNRLCWGLAAISWLGGLGNSVIQSTFTLIQLEPCGHRKVDNFECEVPAM 184

QY      186 LKLACADTSLFDTLTFACCVFMLLPESIIMASYACILGAVLRIRSAQAMKKAATCCSSH 245
      : | | | | | | : : | : : | : | : | : | : | : | : |
Db      185 IKLACGDTSLNEAVLNGVCTFEFTVPVSVILVSCFIAQAVMKIRSVEGRRKAENTCVSH 244

QY      246 LTAVTLEFYGAMEMYLPRRYRAPSHDKVASIFETVLTPLMLNPILYSLRNGEVCALRKRG 305
      | | | | | : : | : | : | : | : | : | : | : | : | :
Db      245 LVVLELFYGSATYGLLPKAKSSNOGKFIISLFYSVTPMVNPPLTYTLRNKEVCALGRL 304

QY      306 LDRCR 310
      | : |
Db      305 LGKGR 309

```

## RESULT 4

probable olfactory receptor OLFR1 - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999  
C;Accession: A48413  
R;Schumans, S.; Muscatelli, F.; Miot, F.; Mattei, M.G.; Vassart, G.; Parmentier, M.  
Cytogenet. Cell Genet. 63, 200-204, 1993  
A;Title: The OLFR1 gene encoding the HGNP07E putative olfactory receptor maps to the  
A;Reference number: A48413; MUID:93251832; PMID:8097991  
A;Accession: A48413  
A;Molecule type: DNA  
A;Residues: 1-312 <SCH>  
A;Cross-references: EMBL:X65857; NID:g425220; PIDN:CAA46687.1; PID:g425221  
A;Note: sequence extracted from NCBI backbone (NCBIN:131483, NCBIIP:131485)  
C;Genetics:  
A;Gene: GDB:OLFR1  
A;Cross-references: GDB:136274; OMIM:164342  
A;Map position: 17p13-17p12  
C;Superfamily: olfactory receptor OR14  
F;26-49/Domain: transmembrane #status predicted <TM1>  
F;57-79/Domain: transmembrane #status predicted <TM2>  
F;101-120/Domain: transmembrane #status predicted <TM3>  
F;140-164/Domain: transmembrane #status predicted <TM4>  
F;197-219/Domain: transmembrane #status predicted <TM5>  
F;237-255/Domain: transmembrane #status predicted <TM6>  
F;272-291/Domain: transmembrane #status predicted <TM7>  
F;5/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F;137,232,290/Binding site: phosphate (Ser) (covalent) #status predicted  
F;312/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match	40.1%;	Score 654.5;	DB 2;	Length 312;
Best Local Similarity	42.28;	Pred. No. 2e-48;		
Matches 128;	Conservative 59;	Mismatches 115;	Indels 1;	Gaps 1;

QY	6	NQSYTDGFFELLGIFSHSQCDLVLFSAVMVFTVALCGNVLLIFLIYLDAGLHTPMYFFLS	65
		:     : : :      :         :	
Db	5	NQSEGEFFLLGMSESPEDQOQLFWMFLSMXLVTVGNVLLIILAISDSRLHTPVYFFELA	64
QY	66	QLSLMDLMLVCNIVPKMAANFLSGRKSISFVGCIGIQIGFVSLVGESEGLLCGLMAYDHYV	125
		:         :    :      :       :	
Db	65	NLSFTDLFFVTNTIPKMLVNLQSHNKAISVAGCLTQLFVLSVALDNLILAAMAYDRYV	124
QY	126	AVSHPLHYPLMNQRVCLQITGSSWAFGLIDGVIQMVAAAGLPRYCGSRSDVHFFCEVOAL	185
		:        : : :   : : :     : :      :     :	
Db	125	AICCPHLHYTTAMSPKLCILILSLCWLVSVLGLIHTLMLRVTFCGSRKRIHYIFCEMYVL	184







```

Db      5 NATEVTDYLLGFGVQONTQCVLIVFVIVYVTSVMGNTGMILLINTNSRLQTPMYFLQ 64
QY      66 QLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGCIQIGFVSLVGEGLLGLMAYDHYV 125
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      65 NLAFVDICYTSAITPKMLQSFMEWEDCSISYTCGIQLLVANFATSDCYLLAVMAVDHYV 124
QY      126 AVSHPLHYPLMNQRCVLCQITGSSWAFGIIDGVIQVAAAMGLPYCGSRSDHFECEVQAL 185
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      125 AICKPLRYPIIMSROVCLLVALSYLMGSINSVHTGFTFSLSYCNKSKNINHEFCDVPI 184
QY      186 LKLACADTSLFDTLFACCVFEMLLPFSTIMASYACILGAVLRIRSAQAMKALATCSSH 245
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      185 ISLSCSNTDINMLLIVFVGFNLTFTVLVITFSYIYIMAILKMSSTAGRKKTFSTCASH 244
QY      246 LTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLNPLIYSLRNGEV 298
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      245 LTAVTIFYGTLSYMYLQPHSDNSEENMKVAVVFGYIVIPMLNPLIYSLRNKEV 297

```

# RESULT 11

S20571  
 Olfactory receptor - dog  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 26-Aug-1999  
 C:Accession: S20571  
 R:Parmentier, M.; Libert, F.; Schurmans, S.; Schiffmann, S.; Lefort, A.; Eggerickx, D.;  
 Nature 355, 453-455, 1992  
 A:Title: Expression of members of the putative olfactory receptor gene family in mammal  
 A:Reference number: S20571; MUID:92131132; PMID:1370859  
 A:Accession: S20571  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-313 <PAR>  
 A:Cross-references: EMBL:X64996; NID:g890; PIDN:CAA46129.1; PID:g891  
 C:Superfamily: olfactory receptor OR14  
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 36.2%; Score 591.5; DB 2; Length 313;  
 Best Local Similarity 39.7%; Pred. No. 4.8e-43;  
 Matches 122; Conservative 62; Mismatches 120; Indels 3; Gaps 3;

```

QY      6 NOSYTDGFFLLGI-FSHSQTDLVLESAMVVFVVALCGNVLLIFLIYLDAGLHTPMYFLL 64
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      5 NOTVSEFVLGLPIDPDQRL-FYALFLAMYVTTLIGNLLIIVLIQLDSHLHTPMYLL 63
QY      65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGCIQIGFVSLVGEGLLGLMAYDHY 124
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      64 SNLSFSDLCFSSVTMPKLLQNMQSQVPSIPYAGCLTQMVFLLFEGDLESFLLVAMAYDRY 123
QY      125 VAVSHPLHYPLMNQRCVLCQITGSSWAFGIIDGVIQVAAAMGLPYCGSRSDHFECEVQA 184
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      124 VAICFPLHYTTIMSPKLCFSLVLVSWVLTFMFAVLHTLMARLCLC-ANTIPHFCDSMA 182
QY      185 LKLACADTSLFDTLFACCVFEMLLPFSTIMASYACILGAVLRIRSAQAMKALATCSS 244
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      183 LKLACSDTQYNELVIFIMGGLIVIPFLITTSYARISSILKVPASIGICKVESTCGS 242
QY      245 HLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLNPLIYSLRNGEVMGALRK 304
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      243 HLSVSLFYGTVIIGLYLCPSSANNSTVKETIAMMYTVTPMLNPFITYSLRNKDMKALRR 302
QY      305 GLDRCRI 311
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      303 VICRKKI 309

```

RESULT 12  
 B23701  
 Olfactory receptor F5 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 04-Sep-1998  
 C:Accession: B23701  
 R:Buck, L.; Axel, R.

Cell 65, 175-187, 1991  
 A:Title: A novel multigene family may encode odorant receptors: a molecular basis for  
 A:Reference number: A23701; MUID:91191556; PMID:1840504  
 A:Accession: B23701  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-313 <BUC>  
 A:Cross-references: GB:M64377  
 C:Superfamily: olfactory receptor OR14  
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 36.2%; Score 591; DB 2; Length 313;  
 Best Local Similarity 38.5%; Pred. No. 5.2e-43;  
 Matches 116; Conservative 56; Mismatches 129; Indels 0; Gaps 0;

```

QY      6 NOSYTDGFFLLGIFSHSQTDLVLESAMVVFVVALCGNVLLIFLIYLDAGLHTPMYFLLS 65
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      5 NQSSVTEFLLGLSRQPQQQLLFLFLIMYIATVNLGILLIILAIIGTDSRLHTPMYFLLS 64
QY      66 QLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGCIQIGFVSLVGEGLLGLMAYDHYV 125
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      65 NLSFVDVCFSSVTTPVKLANHLILGSAISFSGCLTQLYFLAVFGNMDNFLAVMSYDRFV 124
QY      126 AVSHPLHYPLMNQRCVLCQITGSSWAFGIIDGVIQVAAAMGLPYCGSRSDHFECEVQAL 185
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      125 AICPLHYTTKMTKRLQCLVLLVGSWVANMNCILLILMARKSFACADNMIPHFCDGTP 184
QY      186 LKLACADTSLFDTLFACCVFEMLLPFSTIMASYACILGAVLRIRSAQAMKALATCSSH 245
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      185 LKLCSDTHLNDEMILTEGAVVMVTPFVCILISYIHITCAVLRVSSPRGMSFSTCGSH 244
QY      246 LTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIEFTVLTPLNPLIYSLRNGEVMGALRK 305
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      245 LAVCLFYGTVIAYVFNPSSSHLAGRDMAAVMAYAVTTPMLNPFITYSLRNDSMKALRKV 304
QY      306 L 306
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      305 L 305

```

# RESULT 13

JC5200  
 Chemoreceptor TB334 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-Aug-1999  
 C:Accession: JC5200; PC4302  
 R:Thomas, M.B.; Haines, S.L.; Akesson, R.A.  
 Gene 178, 1-5, 1996  
 A:Title: Chemoreceptors expressed in taste, olfactory and male reproductive tissues.  
 A:Reference number: JC5200; MUID:97080538; PMID:8921883  
 A:Accession: JC5200  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-311 <THO1>  
 A:Cross-references: GB:U50947; NID:g1256388; PIDN:AAC52909.1; PID:g1256389  
 A:Accession: PC4302  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 146-153; 265-272 <THO2>  
 A:Experimental source: taste bud  
 C:Comment: This protein is coupled to a GTP-binding protein-mediated signal transduct  
 C:Genetics:  
 A:Gene: tb334  
 C:Superfamily: olfactory receptor OR14  
 C:Keywords: olfaction; taste bud; transmembrane protein  
 F:24-47/Domain: transmembrane #status predicted <TM1>  
 F:56-77/Domain: transmembrane #status predicted <TM2>  
 F:99-118/Domain: transmembrane #status predicted <TM3>  
 F:138-162/Domain: transmembrane #status predicted <TM4>  
 F:195-217/Domain: transmembrane #status predicted <TM5>  
 F:236-258/Domain: transmembrane #status predicted <TM6>  
 F:271-291/Domain: transmembrane #status predicted <TM7>



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:58:31 ; Search time 85 Seconds  
(without alignments)  
763.586 Million cell updates/sec

Title: US-09-634-109D-2  
Perfect score: 1634  
Sequence: 1 MGRWVNSYTDGFLLGIFS.....GEVMGALRKGLDRCRIGSQH 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- SPTREMBL\_21:\*
  - 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriaph:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1440	88.1	315	11	Q8VGD6	Q8vgd6 mus musculu
2	1242	76.0	326	11	Q8VEY7	Q8vey7 mus musculu
3	1044	63.9	223	11	Q9QUL1	Q9qu11 mus musculu
4	1040	63.6	223	11	Q9ZIU2	Q9z1u2 mus musculu
5	1010.5	61.8	317	11	Q8VGD7	Q8vgd7 mus musculu
6	975	59.7	216	4	Q96R30	Q96r30 homo sapien
7	921	56.4	315	11	Q8VF87	Q8vf87 mus musculu
8	893	54.7	313	11	Q8VFA2	Q8vfa2 mus musculu
9	883.5	54.1	317	11	Q8VGT5	Q8vgt5 mus musculu
10	875.5	53.6	309	11	Q8VFG7	Q8vfg7 mus musculu
11	875.5	53.6	312	11	Q8VEY2	Q8vet2 mus musculu
12	858	52.5	312	11	Q8VGX2	Q8vgx2 mus musculu
13	856	52.4	312	11	Q8VEO5	Q8vfo5 mus musculu
14	852	52.1	315	11	Q8VGD8	Q8vgd8 mus musculu
15	828.5	50.7	315	11	Q8VGD9	Q8vgd9 mus musculu
16	815	49.9	310	11	Q8VFD6	Q8vfd6 mus musculu

17	805	49.3	316	11	Q8VEF37	Q8vfe37 mus musculu
18	796	48.7	306	11	Q8VFD5	Q8vfd5 mus musculu
19	793	48.5	308	11	Q8VEFG8	Q8vfg8 mus musculu
20	790	48.3	306	11	Q8VED4	Q8ved4 mus musculu
21	787.5	48.2	312	11	Q8VGJ5	Q8vgj5 mus musculu
22	759.5	46.5	313	11	Q8VGL6	Q8vgl6 mus musculu
23	756	46.3	314	11	Q8VGL7	Q8vgl7 mus musculu
24	742.5	45.4	311	11	Q8VEL3	Q8vel3 mus musculu
25	739.5	45.3	315	11	Q8VEY3	Q8vey3 mus musculu
26	728	44.6	321	11	Q8VEF17	Q8vfe17 mus musculu
27	722	44.2	313	11	Q63394	Q63394 rattus norv
28	716	43.8	321	11	Q8VEF18	Q8vfe18 mus musculu
29	715	43.8	315	11	Q35434	Q35434 rattus norv
30	713	43.6	310	11	Q8VES4	Q8ves4 mus musculu
31	713	43.6	313	11	Q8VEH2	Q8veh2 mus musculu
32	711	43.5	313	11	Q8VES3	Q8ves3 mus musculu
33	710	43.5	310	11	Q8VES5	Q8ves5 mus musculu
34	707	43.3	312	11	Q8VEFNO	Q8vfn0 mus musculu
35	706.5	43.2	314	11	Q9EPG6	Q9epg6 mus musculu
36	706	43.2	315	11	Q8VGB1	Q8vgb1 mus musculu
37	703	43.0	310	11	Q8VGS3	Q8vgs3 mus musculu
38	702	43.0	315	11	Q8VEV9	Q8vev9 mus musculu
39	701	42.9	310	11	Q8VEFS2	Q8vfe2 mus musculu
40	700	42.8	312	11	Q8VEFM0	Q8vfm0 mus musculu
41	698	42.7	312	11	Q8VEF0	Q8vfe0 mus musculu
42	698	42.7	315	11	Q8VEF88	Q8vfe88 mus musculu
43	697.5	42.7	316	11	Q9D3U9	Q9d3u9 mus musculu
44	697	42.7	318	11	Q92022	Q92022 mus musculu
45	696.5	42.6	311	11	Q8VFM6	Q8vfm6 mus musculu

ALIGNMENTS

RESULT 1	Q8VGD6	PRELIMINARY;	PRT;	315 AA.
ID	Q8VGD6:			
AC	Q8VGD6:			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE	Olfactory receptor MOR276-1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Zhang X., Firestein S.J.;			
RT	"The olfactory receptor gene superfamily of the mouse.";			
RL	Nat. Neurosci. 0:0-0(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Adams M.;			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY073214; AAL60877.1; -			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCR_RHODOPSN.			
DR	PROSITE; PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.			
DR	PROSITE; PS50262; G_PROTEIN_RECPE_F1_2; 1.			
KW	Receptor.			
SQ	SEQUENCE 315 AA; 34993 MW; 80A8516DB31846B3 CRC64;			
Qy	Query Match	88.1%;	Score 1440;	DB 11; Length 315;
Db	Best Local Similarity	87.6%;	Pred. No. 3.8e-122;	
Qy	Matches 276; Conservative 16; Mismatches 23; Indels 0; Gaps 0;			
Qy	1 MGRWVNSYTDGFLLGIFSHSQTDLVLSAVVETVALCGNVLLFLIYLDAGLHTPM 60			
Db	1 MGILWNESSVDGFLLGIFSQGTDLLLFSTVMLVETVALCGNVLLIILITDRLHTPM 60			
Qy	61 YFFLSQLMDLMLVCNIVPKMANFLSGRKSTSFVGCQIQIGFFVSLVGSFGLLLGLMA 120			

Db	61	YFFLSQSLMDLMVNCNIVPKMAVNFLSGRKSSISFAGCGIQIGFFVSLVSGEGLLGLMA	1200
QY	121	YDHVVAVSHPLHYPIILMNQRCVLQITGSSWAFGIIDGVQIOWAAMGLPYCGSRSDVDFHC	1800
Db	121	YDRYVAISHPLHYPIILMSQKVCLOIAGSSWAFGILDGIQIOWAAMSLPYCGSRYIDHFFC	1800
QY	181	EVQALLLKLACADTSLFDTLFLACCVFMLLPFSIIMASYACILGAVALRIRSAQAWKKALA	2400
Db	181	EVPALELLKACADTSLFDTLFLACCVFMLLPFSIIVTSARITLGAVALRHHSAQSRKKALA	2400
QY	241	TCSSHLTAVTLFYGAAMFMYLRPRRYRAPS HDKVASIFVTYVLTPLMLNPLIYSLRNGEVMG	3000
Db	241	TCSSHLTAVSLFYGAAMFIYLRPRRYRAPSHDKVVSIFYVLTPLMLNPLIYSLRNRVEMG	3000
QY	301	ALRKGLDRCRIGSQH 315	
Db	301	ALRKGLDRCRIGSQH 315	

```

RESULT 2
08VEY7
ID 08VEY7 PRELIMINARY; PRT; 326 AA.
AC 08VEY7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Olfactory receptor MOR276-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP Zhang X., Firestein S.J.;
RA "The olfactory receptor gene superfamily of the mouse.";
RT Nat. Neurosci. 0:0-0(2002).
RL [2]
RN SEQUENCE FROM N.A.
RP Adams M.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL

```

DR EMBL; AY073756; AAL61419.1; -;  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHDOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; UNKNOWN\_1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 326 AA; 36591 MW; 557FC2F44E528DBF CRC64;

Query Match	76.0%;	Score 1242;	DB 11;	Length 326;
Best Local Similarity	74.8%;	Pred. No. 3e-104;		
Matches 235;	Conservative 30;	Mismatches 49;	Indels 0;	Gaps 0;

[illegible]

```
QY 301 ALRKGLDRCRIGSQ 314
    ||:|||||:||||
Db 312 ALQKGLDRCRVGSQ 325
```

RESULT 3		
09QUL1		
ID 09QUL1	PRELIMINARY;	PRT: 223 AA.
AC 09QUL1;		
DT 01-MAY-2000 (TREMBLrel. 13, Created)		
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE Olfactory receptor G7 (Fragment).		

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL/6; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=99091050; Pubmed=9875846;
RA Krautwurst D., Yau K.W., Reed R.R.;
RT "Identification of ligands for olfactory receptors by functional
RT expression of a receptor library.";
RL Cell 95:917-926(1998).
DR EMBL; AF102537; AAD13329.1; -.
DR EMBL; AF102533; AAD13325.1; -.
DR MGD; MGI:1333785; Olf1r56.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR000237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 223
SQ SEQUENCE 223.AA; 24561 MW; 8D5E772850474A86 CRC64;

```

Query Match	63.9%;	Score 1044;	DB 11;	Length 223;
Best Local Similarity	89.7%;	Pred. No. 1.6e-86;		
Matches 200; Conservative	11;	Mismatches 12;	Indels 0;	Gaps 0;

[illegible]

RESULT 4  
Q92IU2  
ID Q92IU2 PRELIMINARY; PRT; 223 AA

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE Olfactory receptor G3 (Fragment).  
GN OLF56.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.



OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BL/6; TISSUE=OLFACTORY EPITHELIUM;  
RX MEDLINE=99091050; PubMed=9875846;  
RA Krautwurst D., Yau K.W., Reed R.R.;  
RT "Identification of ligands for olfactory receptors by functional  
RT expression of a receptor library."  
RL Cell 95:917-926(1998).  
DR EMBL: AF102535; AAD13327.1; -.  
DR MGD: MGI:1333785; Olfir56.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCRHHODOPSN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 1 1  
FT NON\_TER 223 223  
SQ SEQUENCE 223 AA; 24591 MW; 8D421B5850475F93 CRC64;

Query Match 63.6%; Score 1040; DB 11; Length 223;  
Best Local Similarity 89.2%; Pred. No. 3.6e-86;  
Matches 199; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFFVSLVSGELLLGLMAYDHY 124  
Db 1 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFAGCGIIGFFVSLVSGELLLGLMAYDRY 60  
QY 125 VAVSHPLHYPLIMNQRVCLQITGSSWAFGIIDGVIQVMAAGLPYCGSRVDHFFCEVQA 184  
Db 61 VAISHPLHYPLIMSQKVCLOIAGSSWAFGIIDGIQVMAASLPYCGSRVIDHFFCEVPA 120  
QY 185 LKLLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKALATCSS 244  
Db 121 LKLLACADTSLEFDTLFACCVFMLLPFSITVTSYARILGTVLRHMSAQSRKKALATCSS 180  
QY 245 HLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLN 287  
Db 181 HLTAVSLFYGAAMFYLRPRRYRAPSHDKVVSIFYTVLTPMLN 223

RESULT 5  
Q8VGD7 PRELIMINARY: PRT; 317 AA.  
AC Q8VGD7;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
DE Olfactory receptor MOR277-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY073213; AAL60876.1; -.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCRHHODOPSN.  
DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 317 AA; 35453 MW; 331CD9B8CDCADAB1 CRC64;

Query Match 61.8%; Score 1010.5; DB 11; Length 317;  
Best Local Similarity 61.9%; Pred. No. 2.3e-83;

Matches 192; Conservative 46; Mismatches 71; Indels 1; Gaps 1;  
QY 4 WV-NQSYTDGFELLGIFSHSQTDLVLSAVMVETVALCGNVLLIFLIYLDAGLHTPMYF 62  
Db 3 WAGNOTLISHFVLLGLFTSHPLHLFSLIMVFLVALSGNGLMILLIMDSRLHTPMYF 62  
QY 63 FLSQLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFFVSLVSGELLLGLMAYD 122  
Db 63 FLSWLSLMDLMLISTIVPRNAADFLRGSGISFAGCGLIQIEFTLLGDECFLLAFMAYD 122  
QY 123 HYVAVSHPLHYPLIMNQRVCLQITGSSWAFGIIDGVIQVMAAGLPYCGSRVDHFFCEV 182  
Db 123 RYVAISNPLRSVIMSRRCVWLMVAGSWLEGLVDGLIQAVFTLRFPYCGSQEIDHFFCEV 182  
QY 183 QALLKLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKALATC 242  
Db 183 PAVLKLACADTSLEYETMIYCCVLMMLLPFSIVISASYLRILVAVLRMRSAEGRKAFATC 242  
QY 243 SSHLTAVTLFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTPMLNPLIYSLRNGEVMGAL 302  
Db 243 SSHMIVSLFYGAAMITVMRQAYHSSKQDKVSAFYTMITPMLNPLIYSLRNKEVTGAL 302  
QY 303 RKLGLDRCRIG 312  
Db 303 RKLGLKCPGC 312

RESULT 6  
Q96R30 PRELIMINARY: PRT; 216 AA.  
ID Q96R30;  
AC Q96R30;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE Olfactory receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fuchs T., Malécova B., Linhart C., Sharan R., Khén M., Herwig R.,  
RA Shmulevich D., Elkon R., Steinfath M., O'Brien J.K., Radlof U.,  
RA Lehrach H., Glender Z., Glusman G., Lancet D., Shamir R.;  
RT "DEFOG: A Practical Scheme for Deciphering Families of Genes."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF399614; AAK95099.1; -.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 1 1  
FT NON\_TER 216 216  
SQ SEQUENCE 216 AA; 24042 MW; B7F41AF784E61D20 CRC64;

Query Match 59.7%; Score 975; DB 4; Length 216;  
Best Local Similarity 85.6%; Pred. No. 2.6e-80;  
Matches 185; Conservative 12; Mismatches 19; Indels 0; Gaps 0;  
QY 69 LMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFFVSLVSGELLLGLMAYDHYVANS 128  
Db 1 LMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGLFLVGSSEGLLLGLMAYDRYVAIS 60  
QY 129 HPLHYPLIMNQRVCLQITGSSWAFGIIDGVIQVMAAGLPYCGSRVDHFFCEVQALLKL 188  
Db 61 HPLHYPLIMNQRVCLQITGSSWAFGIIDGIQVMAVVMNFPYCGLRKVNHFCEMLSLKL 120  
QY 189 ACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRIRSAQAMKALATCSSHLTA 248  
Db 121 ACVDTSLEFVKVIFACCVFMLLPFSITVTSYARILGTVLQMHSAQAMKALATCSSHLTA 180  
QY 249 VTLFYGAAMFMYLRPRRYRAPSHDKVASIFYTVLTP 284

Db 181 VTLFYGAAMFYLRPRHYRASHDKVASIFYVLTLP 216

## RESULT 7

Q8VF87 PRELIMINARY: PRT: 315 AA.  
AC Q8VF87;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Olfactory receptor MOR279-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY073647; AAL61310.1; -  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCRHHODOPSN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 315 AA; 35208 MW; 05FB088ACD114099 CRC64;

Query Match 56.4%; Score 921; DB 11; Length 315;  
Best Local Similarity 55.5%; Pred. No. 2.8e-75;  
Matches 172; Conservative 53; Mismatches 83; Indels 2; Gaps 1;

QY 1 MGRVNSQSTYDGFLLGIFSHSOTDLVFSAVMVFTVALCGNVLLIFLYLDAGLHTPM 60  
Db 3 MGRVNSQSTYDGFLLGIFSHSOTDLVFSAVMVFTVALCGNVLLIFLYLDAGLHTPM 60  
QY 61 YFELSLSLMDLMLVCNIVPKMAANFLSGKRSISFVCGGIGFVSVLGSSEGLLGMA 120  
Db 63 YFELSLSLMDLMLVCNIVPKMAANFLSGKRSISFVCGGIGFVSVLGSSEGLLGMA 120  
QY 121 YDHYVAVSHPLHYPIILMNQVCLQITGSSNAFGIIDGVIOVMAAMGLPYCGSRSDHFFC 180  
Db 123 YDHYVAVSHPLHYPIILMNQVCLQITGSSNAFGIIDGVIOVMAAMGLPYCGSRSDHFFC 180  
QY 181 EVQALLKLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRISAQAWKKALA 240  
Db 183 EVQALLKLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRISAQAWKKALA 240  
QY 241 TCSSHLTAATLIFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLNPLIYSLRNGEVMG 300  
Db 243 TCSSHLTAATLIFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLNPLIYSLRNGEVMG 300  
QY 301 ALRKL 306  
Db 303 AFMKVLGIDK 312

## RESULT 8

Q8VFA2 PRELIMINARY: PRT: 313 AA.  
AC Q8VFA2;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Olfactory receptor MOR279-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY073632; AAL61295.1; -  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCRHHODOPSN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 313 AA; 35116 MW; 623926357774A7C0 CRC64;

Query Match 54.7%; Score 893; DB 11; Length 313;  
Best Local Similarity 52.6%; Pred. No. 9.3e-73;  
Matches 161; Conservative 57; Mismatches 88; Indels 0; Gaps 0;

QY 1 MGRVNSQSTYDGFLLGIFSHSOTDLVFSAVMVFTVALCGNVLLIFLYLDAGLHTPM 60  
Db 1 MGRVNSQSTYDGFLLGIFSHSOTDLVFSAVMVFTVALCGNVLLIFLYLDAGLHTPM 60  
QY 61 YFELSLSLMDLMLVCNIVPKMAANFLSGKRSISFVCGGIGFVSVLGSSEGLLGMA 120  
Db 61 YFELSLSLMDLMLVCNIVPKMAANFLSGKRSISFVCGGIGFVSVLGSSEGLLGMA 120  
QY 121 YDHYVAVSHPLHYPIILMNQVCLQITGSSNAFGIIDGVIOVMAAMGLPYCGSRSDHFFC 180  
Db 121 YDHYVAVSHPLHYPIILMNQVCLQITGSSNAFGIIDGVIOVMAAMGLPYCGSRSDHFFC 180  
QY 181 EVQALLKLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRISAQAWKKALA 240  
Db 181 EVQALLKLACADTSLEFDTLFACCVFMLLPFSIIMASYACILGAVLRISAQAWKKALA 240  
QY 241 TCSSHLTAATLIFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLNPLIYSLRNGEVMG 300  
Db 241 TCSSHLTAATLIFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLNPLIYSLRNGEVMG 300  
QY 301 ALRKL 306  
Db 301 AFMKVL 306

## RESULT 9

Q8VGT5 PRELIMINARY: PRT: 317 AA.  
AC Q8VGT5;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Olfactory receptor MOR274-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse."  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY073061; AAL60724.1; -  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.

DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECIP\_F1\_1; UNKNOWN\_1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECIP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 317 AA; 35714 MW; 826485A1D6005E80 CRC64;

Query Match	54.1%;	Score 883.5;	DB 11;	Length 317;
Best Local Similarity	52.8%;	Pred. No. 6.8e-72;		
Matches 161;	Conservative 61;	Mismatches 82;	Indels 1;	Gaps 1;

**OY**

6 NQSYTDGFEFFLLIGFHSQTDLVLFSAVMWVEVTVALCGNVLLIFLIYLDAIGHTPMYEFLS 65  
| | | | : |: : : || | | : ::|||: | | ||||| ||  
**Dd**

5 NTSSD-FTFMGLNTEETSGLVFAATISVFILALVANGIMFIHTDAHHTPMYEFLS 63

```
QY      66 QELMDLMLVCNIVPKMAANFISGRKSISEVCGGIIQIGFNVS LVSGELLGLMAYDHYV 12
       || :| :| : | |||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      64 HLSFTDMYISTVPKMLVDYLIGORTISFVGCTAQHFLYLLTVGAEEFFLLGLMAYDRYV 12
```

QY	126	AVSHPLHYPIILNQVCLQITGSSWAFGLIDGVIQWAAAGLPRYCGSRVDHFFCEVQAL	18
	124	ALCNPLRYPLVLSNRICWIIAGSWGSLDGLLTPITMSFPFCRSREINHFCEAPAV	18

QY 186 LKACADTSLFDTLLFACCVEMLLPFSIMASYACILGANLIRSAQAKALATCSSH 24  
|||||||:::||: ||| |||:::||: ||| | : | : ||| |||||

Ddb 184 LKACADTALYETVMVVCVCYMLLPFSIVTSYARILATVYMSSVEGRKAATCSSH 24

QY 246 LFAVTLFYGAAMMYLRPRRYAPSHDKVASIFVTLTPLMNPILYSLRNGEVMGALRK 30  
:  
Db 244 MIVTTLFYGAALTYTMVPHSYHSPSODKTFSEVYTTITPMTNDITYSMRNDVSGGI RRA 30

QY	306	LDRCR	310
		:	
Dh	304	IGKTP	308

RESULT 10  
08AVEC7

ID	Q8VFG/	PRELIMINARY;	PRI;	309 AA.
AC	Q8VFG/			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Fast ScienceC web44to)			

01 Olfactory receptor MOR278-1.  
02 Mus musculus (Mouse).  
03 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
04

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
BP SEQUENCE FROM N A

RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 0:0-0(2002).  
PM 121

RP SEQUENCE FROM N.A.  
RA Adams M.:  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

```
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR RHODOPSIN.
```

```
DR_PROSITE, PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 309 AA; 34334 MW; 590A86E22A675F82 CRC64;
```

Query Match	53.6%	Score 875.5;	DB 11;	Length 309;
Best Local Similarity	54.1%;	Pred. No. 3.5e-71;		
Matches 164; Conservative	52;	Mismatches 86;	Indels 1;	Gaps

```
QY      6  NQSYTDGFEFLLGIFGSHSQTDLVFSAWVVVFTVALCGNVLLIFLIYDAGLHTPMYFELS 65
      | : | | | | | | : : | | | | | : | | | | | | | | | | | | | | | |
Db      6  NSIEAD-FILLGLFSHTAHSLLSVVLVFTASLMGNITMILLICRDPRLHTPMYFELS 64
```

QY 66 QLSLMDLMLVNCNIVPKMAANFLSGRKSTISFVCGCIQIGFEFVSLVSGEGLLGLMAYDHYV 125  
 |||||:| | |||||:| :||| ||| | :||| |||  
 Db 65 QLSLMDMLVSTIVPKMAANLITSTRSTSPAGCGSQFLFLTLAGECFLLAAMSYDRIV 124

```
QY 126 AVSHPLHYPIIMNQVCLQITGSSWAFGIIDGVIQWVAAMGLPYGSGRSVDHFFCEVOAL 185
    | : | | : | | : | : | | : | : | : | : | : |
Db 125 AICFPLRHYHVMNPKLCALYLVGSWLGCADGLMQAGTILSFPPCHSRRTINHFCEAPSL 184
```

```

Qy      186 EKACADTSLFDTLFACQVEMLLPESIMASYACILGAVLRISQAAMKKAATCSSH 245
      ::||||| :|: :: ||: ||| : ::|||: || |||:| | ||| ||||
Db      185 VRLACADTKVEEFMYICCLIMLLIPSLVLA SYSLIATVLRMSSARKKAFTYCSSH 244

```

```

Qy      246 LTAVTLEFGAAMFMYLRPRRYRAPSHDKVASIFYVLTPLMLNPILSYLRNGEYWGALRKG 305
      | | | | | : | : | : | : | | | | | : | : | : | : | | | | |
Db      245 LAVVGLFYGAITFIYMRPSHQPKSKDVASAFYITFTPLVNLPLIYSVRNKEYGKALRKW 304

```

QY	306	LDR	308
		:	
Db	305	LQK	307

## RESULT 11

ID	NAME	FILE	SIZE	DATE
AC	Q8VET2;			
DT	01-MAR-2002 (Tremblrel, 20, Created)			
DT	01-MAR-2002 (Tremblrel, 20, last sequence update)			

OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:  
 OS Mus musculus (Mouse).  
 DE Olfactory receptor MOR280-1.  
 DI 01.000.1000 (Hirshauer, 21, last annotation update)

UC Mammalia; Eulineid; Rodentia; Sciurognathii; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N A

RA Zindang A., Firestein S.U.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 0:0-0(2002).  
PN 121

RP SEQUENCE FROM N.A.  
RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY073817. AAF61480.1. -

DR InterPro: IPR000276; GPCR\_khodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_RHODPSN.  
DR PROSITE: PS00237; G PROTEIN RECD E1 1. ENZYGM 1

```
DR PROSITE, PS50262; G_PROTEIN_RECEP_F1_2; 1.  
KW Receptor.  
SQ SEQUENCE 312 AA; 34411 MW; 74F1D05EB676307 CRC64;
```

Query Match	53.6%;	Score 875.5;	DB 11;	Length 312;
Best Local Similarity	55.1%;	Pred. NO. 3.5e-71;		
Matches 168; Conservative	44;	Mismatches 92;	Indels 1;	Gaps

QY 4 WNQSYTDGFEFLLGIFSHQTDVLFEASVAVVFYALCGNVLLFLIYLDAGLHTPMYFF 63  
| | | : | | | | : : . : | : | : : : | | | | |  
Db 6 W-NHSSLSFTLAGLFGHSYPDSEFFSLVLLAFGA VGVNILLMTVIQVDRRLHTPMYFF 64

```
QY      64 LSQSLMDLMLVCNIVPKMAANFLSGRKSISFVGCCIQIGFFVSLVGSSEGLLGLMAYDH 123
        |||||:: : | :||||| ||||| || ||||| :: :| :| :|||
Db      65 LSQSLIMDLMTCTVVPKMATNFLSGKLLISLGCSAQIFVVTVGGAECCFLAVMAYDR 124
```

Qy 124 YAVSHPLHLPIILMNORVCLQITGSSNAFGIIDYQIOMVAAMGLPYCGSRSDHFFCEVQ 183  
|::||:||::||::||::||::||::||::||::||::||::||::||::||::||::||  
Db 125 YMAVCYPLRLFPVLMMNKACSFATASNMGMADSVIDGVVFSEFPCGSLEVDHFCEVR 184

QY 184 ALKLIACADTSLEFDTLTACCVMFLLPESIIMASYACILGAVLRIRSAQAAMKKALATCS 243  
| | | | | : | | | | | : | | | | : | | : | : | | |  
Db 185 ALLRLSCADTSLEFDLTYACCVMLLLPLGVIVASARVLTYYWRMPSTEGQKALTTCSS 244

```
Qy      244 SHLTAVTLEYGAAEMFKRPRRYRAPSHDKVASIFVTTLTPMLNPLIYSLRNGEVMGALR 303
      ||| | :||| |:: | | : |||||:||| ||||| || ||:
```

5

Db 245 SHLAVVGLYGAIFSXYQRASARTPLGDRATSIFVTITVPMFNPLIYSLRNREVTSAUK 304  
QY 304 KGLDR 308  
Db 305 KMLER 309

RESULT 12  
Q8VGX2

ID Q8VGX2 PRELIMINARY; PRT; 312 AA.

AC Q8VGX2; 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Olfactory receptor MOR270-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

RN [1]  
RP SEQUENCE FROM N.A.

RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]

RN [2]  
RP SEQUENCE FROM N.A.

RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY073019; AAL60682.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECCEP\_F1\_1; UNKNOWN\_1.

DR PROSITE; PS50262; G\_PROTEIN\_RECCEP\_F1\_2; 1.  
KW Receptor.

SO SEQUENCE 312 AA; 35391 MW; 62504A0407AFC976 CRC64;

Query Match 52.5%; Score 858; DB 11; Length 312;

Best Local Similarity 54.2%; Pred. No. 1.3e-69;

Matches 166; Conservative 55; Mismatches 83; Indels 2; Gaps 2;

QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVFSAVMVFTVALCGNVLLIFLIYLDAGLHTPM 60

Db 1 MEKW-NQSSD-FTLLGLLPQNOTGLLLMLLIIFVESLALCGNSGMHLIRVDPRLHTPM 58

QY 61 YFLLSQSLMDLMLVCNIVPKMANFLSGRKSISFVCGCIQIGFVSLVSGELLLGLMA 120

Db 59 YFLLSQSLMDLMLYSTVTPKMAFNFLSGQKISIFLCGVQSFFFLTMACSEGLLLASMA 118

QY 121 YDHVAVSHPLHYPIILMNQRCVLCQTGSSWAFGIIDGVIQMVAAAGLPYCGSRVDHFFC 180

Db 119 YDRFVAICHPLHYPIIRMSKIMCLKMIGSWILGINSLSLAHTYALHIPYCHSRISINHEFC 178

QY 181 EVQALKLACADTSLFDTLFLACCVEMLLPFSTIMASYACILGAVLRIRSAQAMKKALA 240

Db 179 DVPAMLPLACMDTWVEYEMVFTSLFLLPFLGITASYGRVLFVAFHMRSGEKKKRAFT 238

QY 241 TCSSHLTAVTLFYGAMFMYLRPRRYRAPS HDKVASIFVTYLTPLMLNPLIYSLRNGEYMG 300

Db 239 TCSTHLTVTFYTAPEYTYTLRPRSLRSPTEDEKILAVFYTILTPMLNPLIYSLRNEVIG 298

QY 301 ALRKGL 306

Db 299 AMTRVL 304

## RESULT 13

ID Q8VF05 PRELIMINARY; PRT; 312 AA.

AC Q8VF05; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Olfactory receptor MOR271-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]  
RP SEQUENCE FROM N.A.

RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]

RN [2]  
RP SEQUENCE FROM N.A.

RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY073736; AAL61399.1; -.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCR\_RHODOPSN.

DR PROSITE; PS00237; G\_PROTEIN\_RECCEP\_F1\_1; UNKNOWN\_1.

DR PROSITE; PS50262; G\_PROTEIN\_RECCEP\_F1\_2; 1.  
KW Receptor.

SO SEQUENCE 312 AA; 35536 MW; 2F11DC131EDE829D CRC64;

Query Match 52.4%; Score 856; DB 11; Length 312;

Best Local Similarity 53.6%; Pred. No. 2e-69;

Matches 164; Conservative 57; Mismatches 83; Indels 2; Gaps 2;

QY 1 MGRWVNSYTDGFFLLGIFSHSQTDLVFSAVMVFTVALCGNVLLIFLIYLDAGLHTPM 60

Db 1 MEKW-NQSSD-FITLLGLLPQNOTGLLLMLLIIFVESLALFGNSAMHLIRVDPRLHTPM 58

QY 61 YFLLSQSLMDLMLVCNIVPKMANFLSGRKSISFVCGCIQIGFVSLVSGELLLGLMA 120

Db 59 YFLLSQSLMDLMLYSTVTPKMAFNFLSGQKNISIFLCGVQSFFFLTMACSEGLLLASMA 118

QY 121 YDHVAVSHPLHYPIILMNQRCVLCQTGSSWAFGIIDGVIQMVAAAGLPYCGSRVDHFFC 180

Db 119 YDRFVAICHPLHYPIIRMSKIMCLKMIGSWILGINSLSLAHSIYALHIPYCHSRISINHEFC 178

QY 181 EVQALKLACADTSLFDTLFLACCVEMLLPFSTIMASYACILGAVLRIRSAQAMKKALA 240

Db 179 DVPAMLPLACMDTWVEYEMVFTSLFLLPFLGITASYGRVLFVAFHMRSGEKKKRAFT 238

QY 241 TCSSHLTAVTLFYGAMFMYLRPRRYRAPS HDKVASIFVTYLTPLMLNPLIYSLRNGEYMG 300

Db 239 TCSTHLTVTFYTAPEYTYTLRPRSLRSPTEDEKILAVFYTILTPMLNPLIYSLRNEVIG 298

QY 301 ALRKGL 306

Db 299 AMTRVL 304

RESULT 14

ID Q8VGD8 PRELIMINARY; PRT; 315 AA.

AC Q8VGD8; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Olfactory receptor MOR281-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

RN [1]  
RP SEQUENCE FROM N.A.

RA Zhang X., Firestein S.J.;  
RT "The olfactory receptor gene superfamily of the mouse.";  
RL Nat. Neurosci. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.

RA Adams M.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY073212; AAL60875.1; -.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:59:46 ; Search time 14 Seconds  
(without alignments)  
662.016 Million cell updates/sec

Title: US-09-634-109D-2  
Perfect score: 1634  
Sequence: 1 MGRWVNSQSYTDGFFLLGIFS.....GEVWGALRKGLDRCRIGSOH 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	869	53.2	316	2	US-08-827-291A-2	Sequence 2, Appli
2	636.5	39.0	321	4	US-08-748-506-20	Sequence 20, Appl
3	635.5	38.9	321	4	US-08-748-506-18	Sequence 18, Appl
4	633.5	38.8	321	4	US-08-748-506-12	Sequence 12, Appl
5	632.5	38.7	321	4	US-08-748-506-10	Sequence 10, Appl
6	630.5	38.6	321	4	US-08-748-506-11	Sequence 11, Appl
7	630.5	38.6	321	4	US-08-748-506-13	Sequence 13, Appl
8	630.5	38.6	321	4	US-08-748-506-19	Sequence 19, Appl
9	610.5	37.4	296	2	US-08-467-948A-2	Sequence 2, Appli
10	610.5	37.4	296	3	US-08-467-947A-2	Sequence 5, Appli
11	588	36.0	309	3	US-08-988-876-5	Sequence 6, Appli
12	571	34.9	333	3	US-08-988-876-6	Sequence 7, Appli
13	570	34.9	314	3	US-08-988-876-7	Sequence 61, Appl
14	534	32.7	284	1	PCT-US93-08528-61	Sequence 61, Appl
15	534	32.7	284	5	US-08-467-947A-27	Sequence 27, Appl
16	533	32.6	222	2	US-08-467-947A-27	Sequence 27, Appl
17	533	32.6	222	3	US-08-467-947A-27	Sequence 27, Appl
18	529.5	32.4	286	1	PCT-US93-08528-65	Sequence 65, Appl
19	529.5	32.4	286	5	US-08-748-506-22	Sequence 22, Appl
20	501	30.7	327	4	US-08-748-506-23	Sequence 23, Appl
21	501	30.6	284	1	PCT-US93-08528-67	Sequence 67, Appl
22	499.5	30.6	284	5	US-08-118-270-60	Sequence 60, Appl
23	497.5	30.4	293	1	PCT-US93-08528-60	Sequence 60, Appl
24	497.5	30.4	293	5	US-08-748-506-14	Sequence 14, Appl
25	480	29.4	327	4	US-08-748-506-24	Sequence 24, Appl
26	479	29.3	327	4		

28	468.5	28.7	274	1	US-08-118-270-69	Sequence 69, Appl
29	468.5	28.7	274	5	PCT-US93-08528-69	Sequence 69, Appl
30	458.5	28.1	273	1	US-08-118-270-63	Sequence 63, Appl
31	458.5	28.1	273	5	PCT-US93-08528-63	Sequence 3, Appli
32	457	28.0	247	1	US-08-465-980-3	Sequence 3, Appli
33	457	28.0	247	2	US-09-053-303-3	Sequence 3, Appli
34	457	28.0	247	4	US-09-339-115-3	Sequence 3, Appli
35	457	28.0	247	5	PCT-US95-07093-3	Sequence 68, Appl
36	447.5	27.4	277	1	US-08-118-270-68	Sequence 68, Appl
37	447.5	27.4	277	5	PCT-US93-08528-68	Sequence 62, Appl
38	445	27.2	277	1	US-08-118-270-62	Sequence 62, Appl
39	445	27.2	277	5	PCT-US93-08528-62	Sequence 64, Appl
40	443	27.1	269	1	US-08-118-270-64	Sequence 64, Appl
41	443	27.1	269	5	PCT-US93-08528-64	Sequence 66, Appl
42	433	26.5	275	1	US-08-118-270-66	Sequence 66, Appl
43	433	26.5	275	5	PCT-US93-08528-66	Sequence 2, Appli
44	347	21.2	320	1	US-08-465-980-2	
45	347	21.2	320	2	US-09-053-303-2	

ALIGNMENTS

RESULT 1  
US-08-827-291A-2  
; Sequence 2, Application US/08827291A  
; Patent No. 5874243  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Sathe, Ganesh  
; TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: PA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/827,291A  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 316 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-827-291A-2

Query Match 53.2%; Score 869; DB 2; Length 316;  
Best Local Similarity 54.2%; Pred. No. 2.1e-72;  
Matches 162; Conservative 47; Mismatches 90; Indels 0; Gaps 0;  
QY 6 NOSYTDGFFLLGIFSHSQTDLVLFSAVMVETVALCGNVLLIFLIYDAGLHTPMVFLLS 65













```
Qy 246 LTAVTLEYGAAMFYLRPRRYRAPHSHDKVASIEYTVLTPLMLNPLISLRNGEVMGAL 302
    | : | | | : : | | : | | : | | | | | | | | | : | | |
Db 245 LSVLSLEYCTSLGCVYLLSSAPQSTHTTSSAVMYTVPYPMLNPFYISLRNKDIKAL 301
```

## RESULT 12

```

US-08-988-876-6
; Sequence 6, Application US/08988876
; Patent No. 6063596
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: WITH IMMUNE RESPONSE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,876
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0441 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TEXT:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 205814
US-08-988-876-6

```

Query Match	34.9%;	Score 571;	DB 3;	Length 333;
Best Local Similarity	36.2%;	Pred. No. 5.5e-45;		
Matches 109;	Conservative 61;	Mismatches 131;	Indels 0;	Gaps 0;

QY 6 NOSYTDGFEFLIGFSHSQTDLVLFESAAMVFTVALCGNVLIFELIYLDAGLHPTMYFFLS 65  
 11:: 1 111 : :: : : : : : 1 1 111111111  
 Db 5 NRTRVSEFLLIGFEVENKDLQPLIYGLFELSMYLVTVIGNISIIAIIISDPCLHPTMYFFLS 64  
 QY 66 QLSMDLMLVCNIVPKMAANFLSGRSISFVGCQIQIGFEVSLGSEGLLIGIMAYDHYV 125  
 11:: : : 111 1 : : : : 11 11 : : 11111111  
 Db 65 NISFVDICFISTVTPKMLVNIQTQNNVITYAGCITQIYFELLFVELDNFLTITMAYDRYV 124  
 QY 126 AVSHPLHYPLIMNQRCVQITGSSWAGFIIDVGIVMAAMGEPYCGSRSDVDFCEVQAL 185  
 11:: 1111 : : : : 11 : : : : 1111 : : 11111111  
 Db 125 AICHPMHYTVIMNYKLGFLVLVSWIVSVLHALFQSLMMLALPECTHLEIPHYCEPNQV 184  
 QY 186 LKLACADTSLFDTLFACQVFMLLLPESTIMASYACILGAVLRIRSAQAWKALATCSSH 245

[illegible]

QY	306	L	306
Db	305	L	305

## RESULT 13

```

US-08-988-876-7
; Sequence 7, Application US/08988876
; Patent No. 6063596
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
; TITLE OF INVENTION: WITH IMMUNE RESPONSE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,876
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0441 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 32086
US-08-988-876-7

```

Query Match	34.9%;	Score 570;	DB 3;	Length 314;
Best Local Similarity	38.5%;	Pred. No. 6.4e-45;		
Matches 117; Conservative	60;	Mismatches 125;	Indels 2;	Gaps 1;

```
QY      1 MGRWVNOSYTDEGFELLGTFESHQSOTDLEFSAMVVFTVALCGNVELLFLIYDAGLHTPM   60
        || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2 MGO--NOTSISDFLLGLPIQEQQNCYALFLAMYLTTLTGNNLLIVLIRDSHLHTPM   59
QY      61 YEFLSQLSLMDMLVCNIYPKMANFLSGRKSISFVGCGIQIGFFVSLVGSBGLLGMA   120
        ||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
```

```
QY      121 YDHYAVSHPLHYFILMNQRVCLQITGSSWAFGIIDVYIQMVAAMGLEPYCGSRVDHFFC 180
        ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      120 YDRYAICFPLHYTAIMSPMLCLALVALSWVLTFHAMLHTLLMARLCFCADNVIPHFEC 179
        ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY      181 EVQALLKLCACADTSLEFDTLTFACCVFMILLPESTIMASYACILGAVALRIRSAQAWKKALA 240
        :||| | :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      180 DMSALLKLAFSDTRVNEMWVIFIMGGLIVLPFLLIIGSVARIVSSILKPSSKGICKAFS 239
        :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY      241 TCSHLTAVTLFYGAAEMFYLRPRRYRAPSHDKVASIFTVLTPLMLNDIYSLRNGEVMG 300
        ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      240 TCGSHLSVSLEYGVIVIGLYLCCSSANSSTLKDTVMAMMYTVTPMLNDFIYSLRNDDMG 299
        ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY      301 ALRK 304
        ||| :
Db      300 ALSR 303
```

RESULT 14  
US-08-118-270-61  
; Sequence 61, Application US/08118270

APPLICANT: Murphy, Randall B.  
 APPLICANT: Schuster, David I.  
 TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
 TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
 NUMBER OF SEQUENCES: 348  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BROWDY AND NEIMARK  
 STREET: 419 Seventh Street, N.W., Suite 300  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/118,270  
 FILING DATE: 09-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/943,236  
 FILING DATE: 10-SEP-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Townsend, Kevin G.  
 REGISTRATION NUMBER: 34,033  
 REFERENCE/DOCKET NUMBER: MURPHY=2A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-628-5197  
 TELEFAX: 202-737-3528  
 TELEX: 248633  
 INFORMATION FOR SEQ ID NO: 61:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 284 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-118-270-61

Query Match	32.78;	Score 534;	DB 1;	Length 284;
Best Local Similarity	37.98;	Pred. No. 1.2e-41;		
Matches 106;	Conservative 57;	Mismatches 113;	Indels 4;	Gaps 21

QY 27 VIFSAMVVEFTVALCGNVLLIFLIYLDAGLHTPMKFFELSQLSMDMLVCNIVPKMAANF 86  
 :||::: :||:| | :||||||| | :| :||  
 Db 1 LFLFLFLIMYLATVGLNLIILAIGDSRLHTPMKFFELSNLSFVDVCFSSSTVPKVLANH 60  
 QY 87 LSGRKSISFVCGCIQIGFFVSLVSGEGLLGLMAYDHVAVSHPLHYPILMNGRVCLQIT 146  
 :|:|||| | |:| :|:|||| |||| : :||:

[illegible]

RESULT 15  
PCT-US93-08528-61

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	32.7%;	Score 534;	DB 5;	Length 284;
Best Local Similarity	37.9%;	Pred. No. 1.2e-41;		
Matches 106;	Conservative 57;	Mismatches 113;	Indels 4;	Gaps 2

```
QY    27 VLFSAMVVFVTALCGANVLFLIYLIDAGLHTPMVEFSLQSLSMDLMVCNIYPKMANF   86
      :: ::: : ||::| | | : ||||| | | : | | | | | | | | | | | | | | | |
Db     1 LFLLELLMLATLVGLNLIILAIIGDSRLHTPMYFFLSNLSEVDYCFSSTVPKVLANH   60

QY    87 LSGRKSISEVGCIOIGFEVSVLGSSEGLLGLMAYDHYVAVSHPLHYPIIMNQRYCLOIT   146
      : | : || | | | | : | : | | : | | : | | | | | : : : | : :
Db     61 ILGSOAISPESGCLTQLYFLAVFGMNDFLLAVMSYDRVAICHPLHYTTI--RQLCVLLV   118

QY    147 GSSWAFGLIDGVIQVMAMGLPYCCGRSVDHFCEYQALLKLACADTSLEFTLLFACCVF   206
      || : : : : : : | : ||| : |||::| | | : : :
Db     119 VGSWVANNNCLHLILMARKSFCA--DLPHFECDGTPLLKSCSDTHLNELMITLEGAV   176
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 20:27:35 ; Search time 2661 Seconds

(without alignments)  
10368.080 Million cell updates/sec

Title: US-09-634-109D-1

Perfect score: 948  
Sequence: 1 atgggaagatgggtgaacca.....ggattgacagccagcactga 948

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	948	100.0	948	6	AX087885	AX087885 Sequence
2	948	100.0	1070	6	AX087894	AX087894 Sequence
3	946.4	99.8	183455	2	AC008620	AC008620 Homo sapi
4	935.4	98.7	1192	9	AB065465	AB065465 Homo sapi
5	908.8	95.9	975	6	AX241794	AX241794 Sequence
6	812	85.7	236913	2	AL161615	AL161615 Homo sapi
7	810.4	85.5	948	6	AX375246	AX375246 Sequence
8	810.4	85.5	948	6	AX448545	AX448545 Sequence
9	810.4	85.5	1348	9	AB065675	AB065675 Homo sapi
10	810.4	85.5	183455	2	AC008620	AC008620 Homo sapi
11	809	85.3	945	6	AX241700	AX241700 Sequence
12	807.4	85.2	945	6	AX241793	AX241793 Sequence
13	749.6	79.1	106325	2	AC103024	AC103024 Rattus no
14	743.2	78.4	948	10	AY073214	AY073214 Mus muscu
15	743.2	78.4	123115	10	AL645688	AL645688 Mouse DNA
16	743.2	78.4	170949	2	AL669850	AL669850 Mus muscu
17	679.2	71.6	246828	2	AC109931	AC109931 Rattus no
18	674.4	71.1	981	10	AY073756	AY073756 Mus muscu
19	674.4	71.1	123115	10	AL645688	AL645688 Mouse DNA
20	674.4	71.1	170949	2	AL669850	AL669850 Mus muscu
21	672.4	70.9	965	6	AX305134	AX305134 Sequence
22	610.8	64.4	647	9	AF399482	AF399482 Homo sapi
23	552.4	58.3	650	9	AF399614	AF399614 Homo sapi
24	526.6	55.5	669	10	AF102533	AF102533 Mus muscu
25	526.6	55.5	669	10	AF102537	AF102537 Mus muscu
26	525	55.4	669	10	AF102535	AF102535 Mus muscu
27	442.2	46.6	954	10	AY073213	AY073213 Mus muscu
28	399	42.1	82315	9	AL592313	AL592313 Human DNA
29	390.2	41.2	488	6	AX181424	AX181424 Sequence
30	390.2	41.2	488	9	AF179751	AF179751 Gorilla g
31	389.4	41.1	939	6	AX448869	AX448869 Sequence
32	389.4	41.1	1339	9	AB065952	AB065952 Homo sapi
33	389.4	41.1	159867	9	AL450303	AL450303 Human DNA
34	386.4	40.8	159726	2	AC127949	AC127949 Rattus no
35	386.2	40.7	930	10	AY073559	AY073559 Mus muscu
36	386.2	40.7	936	6	AX448867	AX448867 Sequence
37	386.2	40.7	1336	9	AB065953	AB065953 Homo sapi
38	386.2	40.7	170393	10	AC109605	AC109605 Mus muscu
39	384.6	40.6	73038	2	AC020884_3	Continuation (4 of
40	384.6	40.6	194086	2	AL662903	AL662903 Mus muscu
41	383	40.4	184692	2	AC094035	AC094035 Rattus no
42	383	40.4	195912	2	AC103465	AC103465 Rattus no
43	379.8	40.1	939	6	AX448873	AX448873 Sequence
44	379.8	40.1	1339	9	AB065954	AB065954 Homo sapi
45	379.8	40.1	159867	9	AL450303	AL450303 Human DNA

ALIGNMENTS

RESULT 1  
AX087885  
LOCUS AX087885 948 bp DNA linear PAT 17-MAR-2001  
DEFINITION Sequence 1 from Patent WO0114554.  
ACCESSION AX087885  
VERSION AX087885.1 GI:13396877  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 948)  
AUTHORS Vogeli,G. and Wood,L.S.  
TITLE G protein-coupled receptor expressed in brain  
JOURNAL Patent: WO 0114554-A 1 01-MAR-2001;

```
FEATURES
source
  Location/Qualifiers
    1..948
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      1..948
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAC34662.1"
        /db_xref="GI:13396878"
        /translation="MGRVWVQSYTDGFLLGIFSHSQTDLVLSAVMVFVVALGNV
        LLIFLYDAGLHTPMYFRLSQLSLMDLMLVNCIVPKMANFLSGRKSIFVGCIOI
        GFEVSLVSEGLLGLMAYDHYVAVSHPLHYPILMNQVCLQITGSSWAFGIIDGVIQ
        MVAAMGLPYCGSRSDHFECEVQALLKLACADSLFDLILFACCVFMLLPESIIMAS
        YACILGAVLRIRSAQAMRKALATCSSHLAVTLFYGAAMFMYLRPRRYRAPSHDKVAS
        IFYVLFPMNLPLIYSLRNGEVMGALRKGLDRCRIGSQH"
BASE COUNT      172 a      278 c      230 g      268 t
ORIGIN
Query Match      100.0%; Score 948; DB 6; Length 948;
Best Local Similarity 100.0%; Pred. No. 8.6e-242;
Matches 948; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGATGGGTGAACCAAGTCTACACAGATGGCTTCTCTCTGGGCATCTTTCC 60
Db 1 ATGGAAGATGGGTGAACCAAGTCTACACAGATGGCTTCTCTCTGGGCATCTTTCC 60
QY 61 CACAGCCAGACTGACCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 CACAGCCAGACTGACCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 121 TGTGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 121 TGTGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTAAACATTGTCCA 240
Db 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTAAACATTGTCCA 240
QY 241 AAGATGGCAACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 AAGATGGCAACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 301 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 301 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 361 TATGACCACTACGTGGCGGTTAGCCACCACTTCACTATCCATCTCTCATGAATCAGAGG 420
Db 361 TATGACCACTACGTGGCGGTTAGCCACCACTTCACTATCCATCTCTCATGAATCAGAGG 420
QY 421 GTCTGTCTCCAGATTACTGGGAGCTCCTGGGCTTTGGGATATAGATGAGTGATTCAG 480
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCCTGGGCTTTGGGATATAGATGAGTGATTCAG 480
QY 481 ATGGTGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCACTTTTCTGT 540
Db 481 ATGGTGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCACTTTTCTGT 540
QY 541 GAGGTACAAGCTTTATGAAGCTGGCGTGCAGACACTTCCCTTTTGAACCCCTCTC 600
Db 541 GAGGTACAAGCTTTATGAAGCTGGCGTGCAGACACTTCCCTTTTGAACCCCTCTC 600
QY 601 TTTGCTTGTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 601 TTTGCTTGTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 TGCATCTAGGGGCTGTGCTCGAATACGCTGTCTCAGGCTTGAAAAAGCCCTGGCC 720
Db 661 TGCATCTAGGGGCTGTGCTCGAATACGCTGTCTCAGGCTTGAAAAAGCCCTGGCC 720
QY 721 ACCTGCTCTCCACCTAACAGCTGTACACCTCTTATAGGGGAGCCATGTTATGTAC 780
Db 721 ACCTGCTCTCCACCTAACAGCTGTACACCTCTTATAGGGGAGCCATGTTATGTAC 780
```

```
Db 721 ACCTGCTCTCCACCTAACAGCTGTACACCTCTTATAGGGGAGCCATGTTATGTAC 780
QY 781 CTGAGGCTTAGGCGCTTACCGGGCCCTAGCCATGACAGGTGGCTTATCTTCTACACA 840
Db 781 CTGAGGCTTAGGCGCTTACCGGGCCCTAGCCATGACAGGTGGCTTATCTTCTACACA 840
QY 841 GTCTTACTCCCATGTCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGG 900
Db 841 GTCTTACTCCCATGTCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGG 900
QY 901 GCACGTAGAGAGGGGCTGACCGCTGACAGGATTGGCAGCCAGCACTGA 948
Db 901 GCACGTAGAGAGGGGCTGACCGCTGACAGGATTGGCAGCCAGCACTGA 948
RESULT 2
AX087894      1070 bp      DNA      linear      PAT 17-MAR-2001
LOCUS
DEFINITION Sequence 10 from Patent WO0114554.
ACCESSION AX087894
VERSION AX087894.1 GI:13396883
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1070)
AUTHORS Vogel, G. and Wood, L.S.
TITLE G protein-coupled receptor expressed in brain
JOURNAL Patent: WO 0114554-A 10 01-MAR-2001;
Pharmacia & Upjohn Company (US)
FEATURES
source
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  1..1070
BASE COUNT      207 a      305 c      254 g      304 t
ORIGIN
Query Match      100.0%; Score 948; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 8.6e-242;
Matches 948; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGATGGGTGAACCAAGTCTACACAGATGGCTTCTCTCTCTCTCTCTCTCTCTCT 60
Db 123 ATGGAAGATGGGTGAACCAAGTCTACACAGATGGCTTCTCTCTCTCTCTCTCTCTCTCT 182
QY 61 CACAGCCAGACTGACCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 183 CACAGCCAGACTGACCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
QY 121 TGTGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 243 TGTGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 302
QY 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTAAACATTGTGCCA 240
Db 303 TACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTAAACATTGTGCCA 362
QY 303 TACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTAAACATTGTGCCA 362
Db 423 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
QY 423 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
Db 483 TATGACCACTACGTGGCGGTTAGCCACCACTTCACTATCCATCTCATGAATCAGAGG 542
QY 542 TATGACCACTACGTGGCGGTTAGCCACCACTTCACTATCCATCTCATGAATCAGAGG 542
Db 543 GTCTGTCTCCAGATTACTGGGAGCTCTGTGGGCTTTGGGATAATAGATGAGTGTAC 602
QY 602 GTCTGTCTCCAGATTACTGGGAGCTCTGTGGGCTTTGGGATAATAGATGAGTGTAC 602
```

[illegible]

RESULT 3  
 AC008620  
 LOCUS  
 DEFINITION  
 AC008620 Homo sapiens chromosome 5 clone CTB-14A14, WORKING DRAFT SEQUENCE,  
 8 ordered pieces.  
 AC008620  
 AC008620.8 GI:19224690  
 HTG: HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS  
 SOURCE  
 Homo sapiens.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 183455)  
 DOE Joint Genome Institute.  
 TITLE  
 Sequencing of Human Chromosome 5  
 JOURNAL  
 Unpublished  
 2 (bases 1 to 183455)  
 REFERENCE  
 DOE Joint Genome Institute.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 3 (bases 1 to 183455)  
 REFERENCE  
 DOE Joint Genome Institute.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Mar 7, 2002 this sequence version replaced gi:13676956.  
 COMMENT  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>  
 -----  
 Project Information  
 Center Project Name: 69594  
 Center clone name: CIT978SKB\_14A14  
 -----  
 Summary Statistics  
 Consensus quality: 174802 bases at least Q40  
 Consensus quality: 180052 bases at least Q30

Consensus quality: 181633 bases at least Q20  
Estimated insert size: 197550; agarose-fp estimation  
Estimated insert size: 182755; sum-of-contigs estimation  
Quality coverage: 6.16 in Q20 bases; agarose-fp estimation  
Quality coverage: 6.66 in Q20 bases; sum-of-contigs estimation  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1	8540:	contig of 8540 bp in length
*	8541	8640: gap of unknown length
*	8641	37607: contig of 28967 bp in length
*	37608	37707: gap of unknown length
*	37708	39853: contig of 2146 bp in length
*	39854	39953: gap of unknown length
*	39954	109099: contig of 69146 bp in length
*	109100	109199: gap of unknown length
*	109200	113256: contig of 4057 bp in length
*	113257	113356: gap of unknown length
*	113357	116278: contig of 2922 bp in length
*	116279	116378: gap of unknown length
*	116379	160370: contig of 43992 bp in length
*	160371	160470: gap of unknown length
*	160471	183455: contig of 22985 bp in length.

FEATURES	
source	Location/Qualifiers
	1. .183455

BASE COUNT	47989	a	44173	c	44560	g	45962	t	771	others
ORIGIN	/clone_lib="CalTech.human BAC library B"									

Query Match	99.8%;	Score 946.4;	DB 2;	Length 183455;
Best Local Similarity	99.9%;	Pred. No. 2.4e-241;		
Matches 947;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	1	ATGGGAAGATGCGCTGAACCAAGTCCTACACAGATGGCTTCTTCCTCTTGCGCATCTTTCC	60
Db	102197	ATGGGAAGATGGGTGAACCAAGTCCTACACAGATGGCTTCTTCCTCTTGCGCATCTTTCC	102256
QY	61	CACAGCCAGACTGACCTTGTGCCTCTCTCTGCAAGTTATGGTGCTTACACAGTGGCCCTC	120
Db	102257	CACAGCCAGACTGACCTTGTGCCTCTCTCTGCAAGTTATGGTGCTTACACAGTGGCCCTC	102316
QY	121	TGTGGGAATGTCTTCCTCATCTTCCTCATCTACCTGGACGCTGCACTTACACCCTCATG	180
Db	102317	TGTGGGAATGTCTTCCTCATCTTCCTCATCTACCTGGACGCTGCACTTACACCCTCATG	102376
QY	181	TACTTCTTCTTCAGCCAGCTCTCCCTCATGAGCCTCATGTGGTGTAAACATTGTGCCA	240
Db	102377	TACTTCTTCTTCAGCCAGCTCTCCCTCATGAGCCTCATGTGGTGTAAACATTGTGCCA	102436
QY	241	AAGATGGCAGCCAACCTTCTGTCTGGCAGGAAGTCCATCTCCTTTGTGGCGTGGGCATA	300
Db	102437	AAGATGGCAGCCAACCTTCTGTCTGGCAGGAAGTCCATCTCCTTTGTGGCGTGGGCATA	102496
QY	301	CAATTTGGCTTTTTTGTCTCTCTTTGTGGATCTGAGGGGCTTTGCTGGACTCATGGCT	360
Db	102497	CAATTTGGCTTTTTTGTCTCTCTTTGTGGATCTGAGGGGCTTTGCTGGACTCATGGCT	102556
QY	361	TATGACCACTACGTGGCCGTTAGCCACCCACTTCACTATCCATCCTCATGATCAGAGG	420
Db	102557	TATGACCACTACGTGGCCGTTAGCCACCCACTTCACTATCCATCCTCATGATCAGAGG	102616
QY	421	GTCGTCTCCAGATTACTGGGAGCTCCTGGGCTTTGGGATTAATAGATGAGCTGATTGAG	480

Db 102617 GTCTGTCTCCAGATTACTGGAGCTCCCTGGCCCTTTGGGATATAGATGAGTGAATTCAG 102676

QY 481 ATGGTGGCAGCCATGCGGCTTACCTACTGTGCTCAAGAGCGTGGATCACTTTTCTGT 540  
|||||  
Db 102677 ATGGTGGCAGCCATGCGGCTTACCTACTGTGCTCAAGAGCGTGGATCACTTTTCTGT 102736

QY 541 GAGGTACAGCTTTATTGAAGCTGGCCTGTGCAGACACTCCCTTTTGCACACCCTCCTC 600  
|||||  
Db 102737 GAGGTACAGCTTTATTGAAGCTGGCCTGTGCAGACACTCCCTTTTGCACACCCTCCTC 102796

QY 601 TTTGCTTGTGTGTCTCATGCTTCTCCCTTCCCTCATCATGAGCCCTCTATGCT 660  
|||||  
Db 102797 TTTGCTTGTGTGTCTCATGCTTCTCCCTTCCCTCATCATGAGCCCTCTATGCT 102856

QY 661 TGCATCTAGGGGCTGTGCTCCGAATAGCCTCTGCTCAGGCCCTGAAAAAGCCCTGGCC 720  
|||||  
Db 102857 TGCATCTAGGGGCTGTGCTCCGAATAGCCTCTGCTCAGGCCCTGAAAAAGCCCTGGCC 102916

QY 721 ACCTGCTCTCCACCTAACAGCTGTACCCCTCTTCTATGGGGCAGCCATGTTCAATGAC 780  
|||||  
Db 102917 ACCTGCTCTCCACCTAACAGCTGTACCCCTCTTCTATGGGGCAGCCATGTTCAATGAC 102976

QY 781 CTGAGGCTAGGCGCTACCGGCGCCCTAGCCATGACAAGGTGGCCCTATCTTCTACACA 840  
|||||  
Db 102977 CTGAGGCTAGGCGCTACCGGCGCCCTAGCCATGACAAGGTGGCCCTATCTTCTACACA 103036

QY 841 GTCTTACTCCCATGCTGACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGG 900  
|||||  
Db 103037 GTCTTACTCCCATGCTGACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGG 103096

QY 901 GCACCTGAGGAAGGGCTGACCGCTGACAGATTGGACGACGACTGA 948  
|||||  
Db 103097 GCACCTGAGGAAGGGCTGACCGCTGACAGATTGGACGACGACTGA 103144

RESULT 4 1192 bp DNA linear PRI 23-JUL-2002  
AB065465 Homo sapiens gene for seven transmembrane helix receptor, complete  
LOCUS AB065465 cds, isolate:CBRC7TM\_28.  
DEFINITION AB065465  
ACCESSION AB065465  
VERSION AB065465.1 GI:21928240  
KEYWORDS  
SOURCE Homo sapiens (isolate:CBRC7TM\_28) DNA.  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Putami,K., Matsumoto,S.,  
Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.  
TITLE Genome-wide discovery and analysis of human seven transmembrane  
JOURNAL helix receptor genes  
REFERENCE Unpublished  
AUTHORS Suwa,M.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research  
AUTHORS Center (CBRC), National Institute of Advanced Industrial Science  
TITLE and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan  
JOURNAL (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/  
Tel:81-3-3599-8080, Fax:81-3-3599-8081)  
COMMENT This sequence is a seven transmembrane helix receptor candidate  
predicted from the whole human genome sequences using our automated  
system that contains programs of gene  
finding(GeneDecoder), sequence search, motif-domain assignment and  
transmembrane helix prediction.  
And the sequence is submitted by the collaborative project between  
[Computational Biology Research Center (CBRC), National Institute  
of Advanced Industrial Science and Technology (AIST)] and [Genome  
Science Division, Research Center for Advanced Science and  
Technology (RCAST), University of Tokyo].  
FEATURES  
Source Location/Qualifiers  
1. .1192

/organism="Homo sapiens"  
/isolate="CBRC7TM\_28"  
/db\_xref="taxon:9606"  
/chromosome="5"  
201.992  
/codon\_start=1  
/evidence=not\_experimental  
/product="seven transmembrane helix receptor"  
/protein\_id="BAC05724.1"  
/db\_xref="GI:21928241"  
/translation="MGRMVNOSYTDGFLLGIFSHSOTDLVLSAVVFTVALGNV  
LILFLYLDAGLHPMYFLLSQLSLMDLMLVCNIVKMAANFLSGRSISFVCGIOI  
GFVSLVSGELGLIMAYDRYVAVSHPLHYPIIMQRVCLQITGSSWAFGIIDGVIG  
MVAAMGLPYCGSRSDHFECEVQALKLACADTSLFDTLFACCVFMLLLPFSIIAS  
YACISRCAPNTLCSGLEKSPGHLILPNSCHPLMGSVHVPEA"  
BASE COUNT 232 a 339 c 285 g 336 t  
ORIGIN  
Query Match 98.7%; Score 935.4; DB 9; Length 1192;  
Best local similarity 99.8%; Pred. No. 2e-238;  
Matches 947; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 ATGGGAAGATGGGTGAACCAAGTCCCTACACAGATGGCTTCTCTTGGGCATCTTTTCC 60  
Db 201 ATGGGAAGATGGGTGAACCAAGTCCCTACACAGATGGCTTCTCTTGGGCATCTTTTCC 260  
QY 61 CACAGCCAGACTGACCTGTTCCTCTCTCTGCAAGTATATGCTGCTTCAAGTGGCCCTC 120  
Db 261 CACAGCCAGACTGACCTGTTCCTCTCTCTGCAAGTATATGCTGCTTCAAGTGGCCCTC 320  
QY 121 TGTGGGAATGCTCCTCTCATCTCTCTCATATCACTGAGCGCTGGACTTGCACACCCCATG 180  
Db 321 TGTGGGAATGCTCCTCTCATCTCTCTCATATCACTGAGCGCTGGACTTGCACACCCCATG 380  
QY 181 TACTTCTCTCAGCCAGCAGCTCTCCCTCATGAGCACTCATGTGCTGTGTAACATTTGCA 240  
Db 381 TACTTCTCTCAGCCAGCAGCTCTCCCTCATGAGCACTCATGTGCTGTGTAACATTTGCA 440  
QY 241 AAGATGGCAGCAACTTCTCTGCTGGCAGAGATCCATCTCTTGTGGGCTGTGGCATA 300  
Db 441 AAGATGGCAGCAACTTCTCTGCTGGCAGAGATCCATCTCTTGTGGGCTGTGGCATA 500  
QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTGTGCTGAGCTCATGGCT 360  
Db 501 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTGTGCTGAGCTCATGGCT 560  
QY 361 TATGACCACCTACGCTGGCGTTAGCCACCCACTTCACTATCCCATCTCATGAATCAGAGG 420  
Db 561 TATGACCACCTACGCTGGCGTTAGCCACCCACTTCACTATCCCATCTCATGAATCAGAGG 620  
QY 421 GTCTGTCTCCAGATTACTGGAGCTCCTGGGCTTTGGGATATAGATGAGTGAATTCAG 480  
Db 621 GTCTGTCTCCAGATTACTGGAGCTCCTGGGCTTTGGGATATAGATGAGTGAATTCAG 680  
QY 481 ATGCTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGAATCATTCTTCTGT 540  
Db 681 ATGCTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGAATCATTCTTCTGT 740  
QY 541 GAGGTACAGCTTTATTGAAGCTGGCCTGTGCAGACACTTCCCTTTTGCACACCCTCCTC 600  
Db 741 GAGGTACAGCTTTATTGAAGCTGGCCTGTGCAGACACTTCCCTTTTGCACACCCTCCTC 800  
QY 601 TTTGCTTGTGTGTCTCATGCTTCTCTCCCTTCTCCATCATGAGCCCTCTATGCT 660  
Db 801 TTTGCTTGTGTGTCTCATGCTTCTCTCCCTTCTCCATCATGAGCCCTCTATGCT 860  
QY 661 TGCATC-CTAGGGGCTGTGCTCCGAATAGCGTCTGCTCAGGCGCTGAAAAAGCCCTGGC 719  
Db 861 TGCATCTTAGGGGCTGTGCTCCGAATAGCGTCTGCTCAGGCGCTGAAAAAGCCCTGGC 920  
QY 720 CACCTGCTCTCCCACTAACAGCTGTACCCCTCTTCTATAGGGCAGCCATGTTCAATGTA 779  
Db 921 CACCTGCTCTCCCACTAACAGCTGTACCCCTCTTCTATAGGGCAGCCATGTTCAATGTA 980

QY 780 CCTGAGGCGCTACGGGCGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACAC 839  
 DB 981 CCTGAGGCGCTACGGGCGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACAC 1040  
 QY 840 AGTCTTACTCCATGCTGACACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 899  
 DB 1041 AGTCTTACTCCATGCTGACACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 1100  
 QY 900 GGCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA 948  
 DB 1101 GGCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA 1149  
 RESULT 5  
 AX241794 975 bp DNA linear PAT 26-SEP-2001  
 LOCUS AX241794  
 DEFINITION Sequence 542 from Patent WO0127158.  
 ACCESSION AX241794  
 VERSION AX241794.1 GI:15798669  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1 (bases 1 to 975)  
 AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and Yanai,I.  
 TITLE Olfactory receptor sequences  
 JOURNAL Patent: WO 0127158-A 542 19-APR-2001;  
 FEATURES  
 source Location/Qualifiers  
 1..975  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="(H38g391 nucleotide)"  
 BASE COUNT 174 a 290 c 237 g 274 t  
 ORIGIN  
 Query Match 95.9%; Score 908.8; DB 6; Length 975;  
 Best Local Similarity 99.5%; Pred. No. 2.4e-231;  
 Matches 943; Conservative 0; Mismatches 2; Indels 3; Gaps 3;  
 QY 1 ATGGGAAGATGGGTGAACAGCTCTACACAGATGGCTTCTCTCTTGGGATCTTTCC 60  
 DB 1 ATGGGAAGATGGGTGAACAGCTCTACACAGATGGCTTCTCTCTTGGGATCTTTCC 60  
 QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGAGTTATGTTGTTCTTCACAGTGGCCCTC 120  
 DB 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGAGTTATGTTGTTCTTCACAGTGGCCCTC 120  
 QY 121 TGTGGGAATGTCTCTCTCATCTCATCTTACCTTGACGCTGCACTTCACACCCCATG 180  
 DB 121 TGTGGGAATGTCTCTCTCATCTCATCTTACCTTGACGCTGCACTTCACACCCCATG 180  
 QY 181 TACTTCTCTCAGCCAGCTCTCTCTCATGAGCTCATGTGTTGTTCTGTAACATTGTGCCA 240  
 DB 181 TACTTCTCTCAGCCAGCTCTCTCTCATGAGCTCATGTGTTGTTCTGTAACATTGTGCCA 240  
 QY 241 AAGATGGCAGCAACTCTCTGTGTGGCAGAGAGTCCATCTCTTTGTGGGCTGTGGCATA 300  
 DB 241 AAGAT-GCAGCCAACTCTCTGTGTGGCA-GAAGTCCATCTCTTTGTGGGCTGTGGCATA 298  
 QY 301 CAAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTGTGGGATCATGGCT 360  
 DB 299 CAAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTGTGGGATCATGGCT 358  
 QY 361 TATGACCACTACGTGGCGTTAGCCACCCACTTCACTATCCATCCTCATGATCAGAGG 420  
 DB 359 TATGACCGCTACGTGGCGGTTAGCCACCCACTTCACTATCCATCCTCATGATCAGAGG 418  
 QY 421 GTCTGTCTCCAGATTACTGGGAGCTCTGGGCGTTTGGGATATAGATGAGTGAATTGAG 480  
 DB 419 GTCTGTCTCCAGATTACTGGGAGCTCTGGGCGTTTGGGATATAGATGAGTGAATTGAG 478

QY 481 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCATTCTTCTGT 540  
 DB 479 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCATTCTTCTGT 538  
 QY 541 GAGGTACAAGCTTTATTTGAAGCTGGGCTGTGCAGACACTTCCCTTTTGTGACACCCCTC 600  
 DB 539 GAGGTACAAGCTTTATTTGAAGCTGGGCTGTGCAGACACTTCCCTTTTGTGACACCCCTC 598  
 QY 601 TTTCCTGTCTGTCTCTCATGCTTCTCTCTCTCTCTCTCTCATCATCATGAGCTCTATGCT 660  
 DB 599 TTTCCTGTCTGTCTCTCATGCTTCTCTCTCTCTCTCTCTCATCATCATGAGCTCTATGCT 658  
 QY 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCGCTGGAAGAGCCCTGGCC 720  
 DB 659 TGCATCTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCGCTGGAAGAGCCCTGGCC 718  
 QY 721 ACCTGCTCTCCACCTTACAGCTGTCAACCTCTCTATGAGGCGACCATGTTCATGTAC 780  
 DB 719 ACCTGCTCTCTCCACCTTACAGCTGTCAACCTCTCTATGAGGCGACCATGTTCATGTAC 777  
 QY 781 CTGAGGCGCTAGGCGCTACCGGCGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACACA 840  
 DB 778 CTGAGGCGCTAGGCGCTACCGGCGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACACA 837  
 QY 841 GTCTTACTCCATGCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGG 900  
 DB 838 GTCTTACTCCATGCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGG 897  
 QY 901 GCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA 948  
 DB 898 GCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA 945

RESULT 6  
 AL161615/c 236913 bp DNA linear HTG 06-DEC-2001  
 LOCUS AL161615  
 DEFINITION Homo sapiens chromosome 13 clone RP11-324K20, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\* 52 unordered pieces.  
 ACCESSION AL161615  
 VERSION AL161615.3 GI:9863602  
 HTG: HTGS\_PHASE1; HTGS\_CANCELLED.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 REFERENCE  
 AUTHORS Burton,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On Aug 21, 2000 this sequence version replaced gi:7327747.  
 COMMENT  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 ----- Project Information  
 Center project name: ba324K20  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 204106 bases at least Q40  
 Consensus quality: 217692 bases at least Q30  
 Consensus quality: 225366 bases at least Q20  
 Insert size: 231813; sum-of-contigs  
 Insert size: 232359; 21.7% error; agarose-fp  
 Quality coverage: 2.25x in Q20 bases; sum-of-contigs Quality  
 coverage: 2.82x in Q20 bases; agarose-fp  
 -----

\* NOTE: This is a 'working draft' sequence. It currently



\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 4991: contig of 4991 bp in length
* 4992 5091: gap of 100 bp
* 5092 7436: contig of 2345 bp in length
* 7437 7536: gap of 100 bp
* 7537 25274: contig of 17738 bp in length
* 25275 25374: gap of 100 bp
* 25375 28650: contig of 3276 bp in length
* 28651 28750: gap of 100 bp
* 28751 31057: contig of 2307 bp in length
* 31058 31157: gap of 100 bp
* 31158 36458: contig of 5301 bp in length
* 36459 36558: gap of 100 bp
* 36559 41612: contig of 5054 bp in length
* 41613 41712: gap of 100 bp
* 41713 44933: contig of 3221 bp in length
* 44934 45033: gap of 100 bp
* 45034 47911: contig of 2878 bp in length
* 47912 48011: gap of 100 bp
* 48012 50422: contig of 2411 bp in length
* 50423 50522: gap of 100 bp
* 50523 54916: contig of 4394 bp in length
* 54917 55016: gap of 100 bp
* 55017 57305: contig of 2289 bp in length
* 57306 57405: gap of 100 bp
* 57406 63866: contig of 6461 bp in length
* 63867 63966: gap of 100 bp
* 63967 67393: contig of 3427 bp in length
* 67394 67493: gap of 100 bp
* 67494 69948: contig of 2455 bp in length
* 69949 70048: gap of 100 bp
* 70049 72222: contig of 2174 bp in length
* 72223 72322: gap of 100 bp
* 72323 74676: contig of 2354 bp in length
* 74677 74776: gap of 100 bp
* 74777 76906: contig of 2130 bp in length
* 76907 77006: gap of 100 bp
* 77007 81224: contig of 4218 bp in length
* 81225 81324: gap of 100 bp
* 81325 85907: contig of 4583 bp in length
* 85908 86007: gap of 100 bp
* 86008 88623: contig of 2616 bp in length
* 88624 88723: gap of 100 bp
* 88724 97798: contig of 9075 bp in length
* 97799 97898: gap of 100 bp
* 97899 103226: contig of 5328 bp in length
* 103227 103326: gap of 100 bp
* 103327 108117: contig of 4791 bp in length
* 108118 108217: gap of 100 bp
* 108218 112719: contig of 4502 bp in length
* 112720 112819: gap of 100 bp
* 112820 132527: contig of 19708 bp in length
* 132528 132627: gap of 100 bp
* 132628 135076: contig of 2449 bp in length
* 135077 135176: gap of 100 bp
* 135177 137721: contig of 2545 bp in length
* 137722 137821: gap of 100 bp
* 137822 139980: contig of 2159 bp in length
* 139981 140080: gap of 100 bp
* 140081 144461: contig of 4381 bp in length
* 144462 144561: gap of 100 bp
* 144562 146775: contig of 2214 bp in length
* 146776 146875: gap of 100 bp
* 146876 150579: contig of 3704 bp in length
* 150580 150679: gap of 100 bp
* 150680 155630: contig of 4951 bp in length
* 155631 155730: gap of 100 bp
```

```
* 155731 158259: contig of 2529 bp in length
* 158260 158359: gap of 100 bp
* 158360 160435: contig of 2076 bp in length
* 160436 160535: gap of 100 bp
* 160536 166007: contig of 5472 bp in length
* 166008 166107: gap of 100 bp
* 166108 169792: contig of 3685 bp in length
* 169793 169892: gap of 100 bp
* 169893 178396: contig of 8504 bp in length
* 178397 178496: gap of 100 bp
* 178497 182026: contig of 3530 bp in length
* 182027 182126: gap of 100 bp
* 182127 184776: contig of 2650 bp in length
* 184777 184876: gap of 100 bp
* 184877 189011: contig of 4135 bp in length
* 189012 189111: gap of 100 bp
* 189112 196337: contig of 7226 bp in length
* 196338 196437: gap of 100 bp
* 196438 199835: contig of 3398 bp in length
* 199836 199935: gap of 100 bp
* 199936 202963: contig of 3028 bp in length
* 202964 203063: gap of 100 bp
* 203064 206573: contig of 3510 bp in length
* 206574 206673: gap of 100 bp
* 206674 209337: contig of 2664 bp in length
* 209338 209437: gap of 100 bp
* 209438 220044: contig of 10607 bp in length
* 220045 220144: gap of 100 bp
* 220145 224133: contig of 3989 bp in length
* 224134 224233: gap of 100 bp
* 224234 226476: contig of 2243 bp in length
* 226477 226576: gap of 100 bp
* 226577 231337: contig of 4761 bp in length
* 231338 231437: gap of 100 bp
* 231438 234031: contig of 2594 bp in length
* 234032 234131: gap of 100 bp
* 234132 236913: contig of 2782 bp in length.
```

## FEATURES

## source

```
1. .236913
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-324K20"
/clone_lib="RPCI-11.2"
1. .4991
/note="assembly_fragment:00518
fragment_chain:1"
5092. .7436
/note="assembly_fragment:00121
fragment_chain:1"
7537. .25274
/note="assembly_fragment:00704
fragment_chain:1"
25375. .28650
/note="assembly_fragment:00589
fragment_chain:1"
28751. .31057
/note="assembly_fragment:01387
fragment_chain:1"
31158. .36458
/note="assembly_fragment:00814
fragment_chain:2"
36559. .41612
/note="assembly_fragment:00002
fragment_chain:2"
41713. .44933
/note="assembly_fragment:00853
fragment_chain:2"
45034. .47911
/note="assembly_fragment:01260
fragment_chain:3"
48012. .50422
/note="assembly_fragment:00867
```

```

misc_feature      fragment_chain:3"
                    50523. .54916
                    /note="assembly_fragment:00443
                    fragment_chain:3"
                    55017. .57305
                    /note="assembly_fragment:00178
                    fragment_chain:4"
                    57406. .63866
                    /note="assembly_fragment:00929
                    fragment_chain:4"
                    63967. .67393
                    /note="assembly_fragment:00184
                    fragment_chain:5"
                    67494. .69948
                    /note="assembly_fragment:00980
                    fragment_chain:5"
                    70049. .72222

```

Query Match 85.7%; Score 812; DB 2; Length 236913;  
Best Local Similarity 91.0%; Pred. No. 1.7e-205;  
Matches 863; Conservative 0; Mismatches 85; Indels 0; Gaps 0

QY	1	ATGGGAAGATGGGTGAACCAAGCTCTACACAGATGGCTTCTTCTCTTGGGCACTTTTCC	60
Db	170963	ATGGAGACGCTGGGTGAACCAAGCTCTACACAGATGGCTTCTTCTCTTAGGCACTTCTCC	170904
QY	61	CACAGCCAGACTGACCTTGTCCTCTTCTCTGCACTTATGGTGTCTTCACAGTGGCCCTC	120
Db	170903	CACAGTACTGCTGACCTTGTCCTCTTCTCCGGTATATGGCGCTTTCACAGTGGCCCTC	170844
QY	121	TGTGGGAATGTCCTCCTCATCTTCTCATCTTACCTGGAGCGCTGAGACTTCACACCCCATG	180
Db	170843	TGTGGGAATGTCCTCCTCATCTTCTCATCTTACATGGACCCCTACCTTCACACCCCATG	170784
QY	181	TACTTCTCTCCAGCCAGCTCTCCCTCATGAGACCTCATGTGTGCTGTAACTTGTGCCA	240
Db	170783	TACTTCTCTCCAGCCAGCTCTCCCTCATGAGACCTCATGTGTGCTGTAACTTGTGCCA	170724
QY	241	AAGATGGCAGCCAACTTCTCTGTGTGGCAGGAATCCATCTCTTGTGGGCTGTGGCATA	300
Db	170723	AAGATGGCAGCCAACTTCTCTGTGTGGCAGGAATCCATCTCTTGTGGGCTGTGGCATA	170664
QY	301	CAAAATTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTCTTGTCTGGGACTGAGCT	360
Db	170663	CAAAATTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTCTTGTCTGGGACTGAGCT	170604
QY	361	TATGACCACTACGTTGGCCGTTAGCCACCACCTTCACTATCCATTCCTCATGATFCAGAGG	420
Db	170603	TATGACCCGATATGTGGCCATTAGCCACCACCTTCACTATCCATTCCTCATGATFCAGAGG	170544
QY	421	GTCGTCTCCAGATTACTGGAGCTCCTGGCCTTTGGATATATAGATGAGTGAATTGAG	480
Db	170543	GTCGTCTCCAGATTACTGGAGCTCCTGGCCTTTGGATATATGATGAGTGAATCCAG	170484
QY	481	ATGCTGGCAGCCATGAGGCTTACTGTGCTCAAGAGCGTGGATCACTTTTCTGT	540
Db	170483	ATGCTGGTAGTAAATGAAATTTCCCTACTGTGGCTTGAGAGAGTGAACCACTTCTTGT	170424
QY	541	GAGGTACCAAGCTTATTTGAAGCTGGCCTGTGCAGACACTTCCCTTTTGAACCCCTCCTC	600
Db	170423	GAGATGCTATCTTGTGAAGCTGGCCTGTGTAGACACATCCCTGTTTGAGAGGTGATA	170364
QY	601	TTTTCCTTGTCTGTCTTCATGCTTCTCCTTCTCCCTTCATCATATGAGCGCTCCTATGCT	660
Db	170363	TTTTCCTTGTCTGTCTTCATGCTTCTTCTCCCATTTCTCATATGCTGTGGCTCCTATGCT	170304
QY	661	TGCATCTTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCGTGGAAAAAGCCCTGGCC	720
Db	170303	CGCATTTCTAGAGCTGTGCTCGAAATGACACTGTGCTCAGGCGTGGAAAAAGCCCTGGCC	170244
QY	721	ACCTGCTCTCCCACTTAACAGCTGTCAACCCTCTCTATATGGGAGCCATGTTTCATGTAC	780
Db	170243	ACCTGCTCTCCCACTTGACAGCTGTCAACCCTCTCTATATGGGAGCCATGTTTCATCTAC	170184

QY	781	CTGAGCGCTAGGGCGCTACCGGGCCCTAGCCATGACAAGGTGGCTCTATCTTCTACACA	840
Db	170183	CTGAGCGCTAGGCACTACCGGGCCCGCCAGCCATGACAAAGTGGCTCTATCTTACACG	170124
QY	841	GTCCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGTGATGGGG	900
Db	170123	GTCCTTACTCCCATGCTCAACCCCTCATTTACAGCTTGAGGAACAGGAGTGATGGGG	170064
QY	901	GCACTGAGGAAGGGGCTGGACCGCTGCAGGATTGGCAGCCAGCACTGA	948
Db	170063	GCACTGAGGAAGGGGCTGGACCGCTGCAGGATGGCAGCCAGCACTGA	170016

RESULT	7
AX375246	
LOCUS	AX375246
DEFINITION	Sequence 36 from Patent WO0210387.
ACCESSION	AX375246
VERSION	AX375246.1
KEYWORDS	GI:19169964
SOURCE	human.

ORGANISM      Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1  
Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A. R.,

Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Walia, N.K., Hafalia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L., Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.

**TITLE** G-Protein coupled receptors  
**JOURNAL** Patent: WO 0210387-A 36 07-FEB-2002  
Incyte Genomics, Inc. (US)

FEATURES	Location/Qualifiers
source	1. .948

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Inocyte ID No: 55036418CB1"
BASE COUNT      176 a      279 c      228 t      265 t

```

ORIGIN			
Query Match	85.5%;	Score 810.4;	DB 6;
Best Local Similarity	90.9%;	Pred. No. 4.2e-205;	Length 948;
Matches 862;	Conservative 0;	Mismatches 86;	Indels 0;
			Gaps 0;

QY 1 ATGGAGATGGGTGACCAGTCTACACAGATGGCTTCTCTTGGCATCTTTCC 60  
||||| 1 ||||||||||||||||||||||||||||||||||| 111

Db 1 ATGAGACGTGGGTGACCAGTCTACACAGATGGCTTCTCTTGAAGCATCTTCTCC 60

Oy 61 CACAGCCAGACTGACCTTGCTCTCTCTCGACAGTTATGTTGCTTCACAGTGCCCTC 120  
||||| |||||||

Db 61 CACAGTACTGCTGACCTTGCTCTCTCTCGATGTTATGAGGGCTCTTCACAGTGCCCTC 120  
||||| |||||||

oy 121 TGTGGAAATGTCCTCCTCAATCTTCTCCTCACTACCTGGACCGTGGACTTCACACCCCATG 180

Db 121 TGTGGAAITGCTCCTCAITCTTCCATCATGACCGCTCACCTTCACACCCCCATG 180

QY 181 TACTTTCCTCAGCCAGCTCTCCCTCATGAGACCTCATGTGGTCTGTAAATTTGTGCCA 240

```
Db      181 TACTTCTCCTCAGCCAGCTCTCCCTCATGACCTCATGTGGTCTGTACCATGTGCCA 240
```

QY 241 AAGATGGACGCCAACTTCCTGTCTGGCAGGAAAGTCCATCTCCTTTGTGGGCTGTGGCATA 300

Db 241 AAGATGGACGCCAACTTCCTGTCTGGCAGGAAAGTCCATCTCCTTTGTGGGCTGTGGCATA 300

QY 301 CAATTGCGTTTTTGTCTCCTTGTTGGATCTGAGGGCTTGTGGACTCATGGCT 360  
||||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 301 CAAATTGGCCTCTTTGTCTGTCTTGTGTGGAACTGAGGGCCTCTGTCTGGCACTCATAGGCT 360

QY 361 TATGACCACTACGTGGCGTTAGCCACCACTTCACCTATGCCATCCTCATGAATCAGAG 420

Db 361 TATGACCGCTATGTGGCATTAGCCACCCACTTCACCTATGCCATCCTCATGAATCAGAGG 420

OY 421 GTCGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGGATATAGATGAGTGTATCAG 480  
|||||  
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGGATATAGATGAGTGTATCAG 480  
OY 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGCATCTTTTCTGT 540  
|||||  
Db 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGCATCTTTTCTGT 540  
OY 541 GAGGTACAGCTTTATGAGCTGGGCTGTGACAGACATTCCTTTTGACACCCCTCTTC 600  
|||||  
Db 541 GAGGTACAGCTTTATGAGCTGGGCTGTGACAGACATTCCTTTTGACACCCCTCTTC 600  
OY 601 TTTGCTGTCTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
|||||  
Db 601 TTTGCTGTCTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
OY 661 TGCATCTAGGGGCTGTGCTCCGAAATACGCTCTGCTCAGGCTTGAAAAAGCCCTGGCC 720  
|||||  
Db 661 CACATTTAGGGGACTGTGCTGCAAAATGCACTCTGCTCAGGCTTGAAAAAGCCCTGGCC 720  
OY 721 ACCGTCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTTCATGAC 780  
|||||  
Db 721 ACCGTCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTTCATGAC 780  
OY 781 CTGAGGCTTAGGGGCTGTACGGGCTTACGCTGCTCAGAGCTGAGGCTTCTCTACACA 840  
|||||  
Db 781 CTGAGGCTTAGGGGACTGTACGGGCTTACGCTGCTCAGAGCTGAGGCTTCTCTACACG 840  
OY 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900  
|||||  
Db 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900  
OY 901 GCACGTAGGAGGGGCTGTGACCGCTGACAGATTGGCAGCCAGCACTGA 948  
|||||  
Db 901 GCACGTAGGAGGGGCTGTGACCGCTGACAGATTGGCAGCCAGCACTGA 948

RESULT 8  
AX448545  
LOCUS AX448545 948 bp DNA linear PAT 03-JUL-2002  
DEFINITION Sequence 211 from Patent WO0224726.  
ACCESSION AX448545  
VERSION AX448545.1 GI:21697443  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Veithen, A.  
TITLE Olfactory and pheromones g-protein coupled receptors  
JOURNAL Patent: WO 0224726-A 211 28-MAR-2002;  
ChemCom S.A. (BE)  
FEATURES  
source Location/Qualifiers  
1. 948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1. 948  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD37591.1"  
/db\_xref="GI:21697444"  
/translation="METWVNSQSYTDGFLIGISHSTADLVLSVMAVFTVALGCV  
LILFLIMDPHLHTPMYFELSQLSLMDLMLVCTNPKMAANFLSGKRSISFVGCIGI  
GLFVCLVSGELLGLIMAYDRVAISHPLHPILMNORVCLQITGSSWAFGIIDGLIO  
MVVVMNFPYCGLRKVNHFCEMLSLKLACVDTSLEKYIFACCVFMLEPFSIIVAS  
YAHILGTVLQMHSAQAQKALATCSSHLTAVTLLFYGAAMFYLPRHYRAPSHDKVAS  
IEFTVLTPLMLNPLIYSLRNEVMGALRKGLDRCRIGSOH"

BASE COUNT 176 a 279 c 228 g 265 t  
ORIGIN  
Query Match 85.5%; Score 810.4; DB 6; Length 948;  
Best Local Similarity 90.9%; Pred. No. 4.2e-205;

Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
OY 1 ATGGAGAGATGGGTGAACCAAGTCTTACACAGATGGCTTCTCTCTCTCTCTCTCT 60  
|||||  
Db 1 ATGGAGAGCTGGGTGAACCAAGTCTTACACAGATGGCTTCTCTCTCTCTCTCTCT 60  
OY 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
|||||  
Db 61 CACAGTACTGCTGACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
OY 121 TGTGGAGATGCT 180  
|||||  
Db 121 TGTGGAGATGCT 180  
OY 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
|||||  
Db 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
OY 241 AAGATGGAGCCCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
|||||  
Db 241 AAGATGGAGCCCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
OY 301 CAAATTTGGCTTTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
|||||  
Db 301 CAAATTTGGCTCTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
OY 361 TATGACCACTAGTGGGCTTACGACCACTTACCTATCTCTCTCTCTCTCTCTCT 420  
|||||  
Db 361 TATGACCGCTATGTGGGCTTACGACCACTTACCTATCTCTCTCTCTCTCTCT 420  
OY 421 GTCTGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGGATATAGATGAGTGTATCAG 480  
|||||  
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGGATATAGATGAGTGTATCAG 480  
OY 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGCATCTTTTCTGT 540  
|||||  
Db 481 ATGGTGGAGTAATGAATTTCCCTACTGTGGCTTGAGGAAGTGAACATTTCTTCTGT 540  
OY 541 GAGGTACAGCTTTATGAGCTGGGCTGTGACAGACATTCCTTTTGACACCCCTCTTC 600  
|||||  
Db 541 GAGATGCTATCTTGTGTAAGCTGGGCTGTGTAAGACATTCCTTTTGAGAAAGTGATGA 600  
OY 601 TTTGCTGTCTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
|||||  
Db 601 TTTGCTGTCTGTCTTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
OY 661 TGCATCTAGGGGCTGTGCTCCGAAATACGCTCTGCTCAGGCTTGAAAAAGCCCTGGCC 720  
|||||  
Db 661 CACATTTAGGGGACTGTGCTGCAAAATGCACTCTGCTCAGGCTTGAAAAAGCCCTGGCC 720  
OY 721 ACCGTCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTTCATGAC 780  
|||||  
Db 721 ACCGTCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTTCATGAC 780  
OY 781 CTGAGGCTTAGGGGCTTACGGGCTTACGCTGATGACAGCTGGCTTCTCTACACA 840  
|||||  
Db 781 CTGAGGCTTAGGGGACTTACGGGCTTACGCTGATGACAGCTGGCTTCTCTACACG 840  
OY 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900  
|||||  
Db 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900  
OY 901 GCACGTAGGAGGGGCTGTGACCGCTGACAGATTGGCAGCCAGCACTGA 948  
|||||  
Db 901 GCACGTAGGAGGGGCTGTGACCGCTGACAGATTGGCAGCCAGCACTGA 948

RESULT 9  
AB065675  
LOCUS AB065675 1348 bp DNA linear PRI 23-JUL-2002  
DEFINITION Homo sapiens gene for seven transmembrane helix receptor, complete  
cds, isolate:CBRC7TM\_238.  
ACCESSION AB065675

VERSION AB065675.1 GI:21928623  
 KEYWORDS  
 SOURCE Homo sapiens (isolate:CBRC7TM\_238) DNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,  
 Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.  
 Genome-wide discovery and analysis of human seven transmembrane  
 helix receptor genes  
 TITLE  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1348)  
 AUTHORS Suwa,M.  
 JOURNAL Direct Submission  
 TITLE Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research  
 Center (CBRC), National Institute of Advanced Industrial Science  
 and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan  
 (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,  
 Tel:81-3-3599-8080, Fax:81-3-3599-8081)  
 COMMENT This sequence is a seven transmembrane helix receptor candidate  
 predicted from the whole human genome sequences using our automated  
 system that contains programs of gene  
 finding(Genedecoder), sequence search, motif-domain assignment and  
 transmembrane helix prediction.  
 And the sequence is submitted by the collaborative project between  
 [Computational Biology Research Center (CBRC), National Institute  
 of Advanced Industrial Science and Technology (AIST)] and [Genome  
 Science Division, Research Center for Advanced Science and  
 Technology (RCAST), University of Tokyo].  
 FEATURES  
 source  
 1..1348  
 /organism="Homo sapiens"  
 /isolate="CBRC7TM\_238"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 201..1148  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="seven transmembrane helix receptor"  
 /protein\_id="BAC05900.1"  
 /db\_xref="GI:21928624"  
 /translation="METWVNSQSYTDGFFLLGIFSHSTADLVLSVMAVFYALCGNV  
 LLIPLIYMDPHLHTPMYFFLSQLSLMDLMLVCTNPKMANFLSGRKSISFVCGIOI  
 GLFVCLVSGEGLLGLMAYDRVAISHPLHPIILMNQVLCIOITGSSWAFGIIDGLIO  
 MYVVMNFPYCGLRKVNHFCEMLSLKLACVDYSLFEKVIIFACCVFMLLPESIIVAS  
 YAHILGTVLQMSAQAMKALATCSSHLTAVTLFYGAMFTYLRPHRYRAPSHDKVAS  
 IFYTVLTPMLNPILYSLRNREVGALRKGLDRCRIGSQH"  
 CDS  
 BASE COUNT 267 a 376 c 310 g 395 t  
 ORIGIN  
 Query Match 85.5%; Score 810.4; DB 9; Length 1348;  
 Best Local Similarity 90.9%; Pred. No. 4.2e-205;  
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
 QY 1 ATGGGAGATGGGTGAACCACTGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 60  
 Db 201 ATGGAGACGTGGGTGAACCACTGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 260  
 QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTGAGTTATGGTGTCTTACAGTGGCCCTC 120  
 Db 261 CACAGTACTGCTGACCTTGTCTCTCTCTGAGTTATGGTGTCTTACAGTGGCCCTC 320  
 QY 121 TGTGGGAATGCTCTCTCATCTTCTCATCTGAGCTGAGCTTGACCTTACACCCCATG 180  
 Db 321 TGTGGGAATGCTCTCTCATCTTCTCATCTGAGCTGAGCTTGACCTTACACCCCATG 380  
 QY 181 TACTTCTTCTCAGCCAGCTTCTCTCATGAGACCTCATGTGCTGTAACTTGTGCCA 240  
 Db 381 TACTTCTTCTCAGCCAGCTTCTCTCATGAGACCTCATGTGCTGTAACTTGTGCCA 440  
 QY 241 AAGATGGACCACTTCTCTGTGGCAGAGATCCATCTCTTGTGGCTGTGGCATA 300  
 Db 301 AAGATGGACCACTTCTCTGTGGCAGAGATCCATCTCTTGTGGCTGTGGCATA 360

Db 441 AAGATGGACCACTTCTCTGTGGCAGAGATCCATCTCTTGTGGCTGTGGCATA 500  
 QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCT 360  
 Db 501 CAAATTTGGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCT 560  
 QY 361 TATGACCACTTACGTGGCCGTAGCCACCACCTTCACTATCCATCCTCATGATCAGAGG 420  
 Db 561 TATGACCGCTATGTGGCCATTAAGCCACCACCTTCACTATCCATCCTCATGATCAGAGG 620  
 QY 421 GTCTGTCTCCAGATTAATGAGAGCTCCTGGGCTTTGGGATATAGATGAGACTGATTCAG 480  
 Db 621 GTCTGTCTCCAGATTAATGAGAGCTCCTGGGCTTTGGGATATAGATGAGACTGATTCAG 680  
 QY 481 ATGTTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGATCATTTTCTGT 540  
 Db 681 ATGTTGGTATGATGATTAATTTCCCTTACTGTGGCTTGAAGAGGTGAACCATTTCTCTGT 740  
 QY 541 GAGGTACAAGCTTTATGAAGCTGGCCTGTGACAGACACTTCCCTTTTGAACCCCTCTC 600  
 Db 741 GAGATGCTATCTTGTGTAAGCTGGCCTGTGTAAGACACATCCCTGTTGAGAAGGTGATA 800  
 QY 601 TTTGCTTGTCTGTCTTCAATCTTCTCTTCCCTTCTCCATCATCATGAGGCTCTATGCT 660  
 Db 801 TTTGCTTGTCTGTCTTCAATCTTCTCTTCCCTTCTCCATCATCATGAGGCTCTATGCT 860  
 QY 661 TGCATCTAGGGGCTGTGCTCCGAATAGCGTCTGCTCAGAGCTGAAAAAGCCCTGGCC 720  
 Db 861 CACATTTAGGGAAGCTGTGCTCCGAATAGCGTCTGCTCAGAGCTGAAAAAGCCCTGGCC 920  
 QY 721 ACCTGCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTCATGTAC 780  
 Db 921 ACCTGCTCTCCACCTTAACAGCTGTACACCTCTTCTATGAGGGGAGCCATGTCATGTAC 980  
 QY 781 CTGAGGCTTAGGGGCTTACCGGGGCCCCTAGCCATGACAGAGGTGGCTTATCTTACACA 840  
 Db 981 CTGAGGCTTAGGGGCTTACCGGGGCCCCTAGCCATGACAGAGGTGGCTTATCTTACACG 1040  
 QY 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
 Db 1041 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 1100  
 QY 901 GCACTGAGGAAGGGGCTTGACCGCTGACAGATTTGGCAGCCAGCACTGA 948  
 Db 1101 GCACTGAGGAAGGGGCTTGACCGCTGACAGATTTGGCAGCCAGCACTGA 1148  
 RESULT 10  
 AC008620/c 183455 bp DNA linear HTG 07-MAR-2002  
 LOCUS Homo sapiens chromosome 5 clone CTB-14A14, WORKING DRAFT SEQUENCE,  
 DEFINITION 8 ordered pieces.  
 ACCESSION AC008620  
 VERSION AC008620.8 GI:19224690  
 KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 183455)  
 DOE Joint Genome Institute.  
 TITLE Sequencing of Human Chromosome 5  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 183455)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 REFERENCE 3 (bases 1 to 183455)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA





Best Local Similarity 91.0%; Pred. No. 9.9e-205;  
Matches 860; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1 ATGGGAAGATGGGTGAACAGTCTTACACAGATGGCTCTCTCTTGGGCATCTTTTC 60  
Db 1 ATGGAGAGCTGGGTGAACAGTCTTACACAGATGGCTCTCTCTTGGGCATCTTTTC 60  
QY 61 CACAGCCAGACTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
Db 61 CACAGTACTGCTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
QY 121 TGTGGGAATGCT 180  
Db 121 TGTGGGAATGCT 180  
QY 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
Db 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
QY 241 AAGATGGCAGCCCACT 300  
Db 241 AAGATGGCAGCCCACT 300  
QY 301 CAAATGGCTTTTGTCT 360  
Db 301 CAAATGGCTTTTGTCT 360  
QY 361 TATGACCACTACGTGGCGCTTAGCCACCCTTCACTATCCATCTCTCTCTCTCTCT 420  
Db 361 TATGACCGCTATGTGGCGCTTAGCCACCCTTCACTATCCATCTCTCTCTCTCTCT 420  
QY 421 GTCTGTCTCCAGATTACTGGGAGCTCCCTGGGCTTTGGGATATAGATGAGTGTAG 480  
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCCCTGGGCTTTGGGATATAGATGAGTGTAG 480  
QY 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGGATCACTTTTCTGT 540  
Db 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGGATCACTTTTCTGT 540  
QY 541 GAGGTACAGCTTATGGAAGCTGGGCTGTGACAGACACTTCCCTTTTGAACACCCCTCTC 600  
Db 541 GAGGTACAGCTTATGGAAGCTGGGCTGTGACAGACACTTCCCTTTTGAAGAGGTGATA 600  
QY 601 TTTGCTGTGTCT 660  
Db 601 TTTGCTGTGTCT 660  
QY 661 TGCATCTAGGGGCTGTGCTCCGAATAGCGCTGTGCTGAGGCTTGAAAAAGCCCTGCGC 720  
Db 661 CGCATTTCTAGGGGCTGTGCTCCGAATAGCGCTGTGCTGAGGCTTGAAAAAGCCCTGCGC 720  
QY 721 ACCTGCTCTCCACCTTACAGCTGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
Db 721 ACCTGCTCTCCACCTTACAGCTGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
QY 781 CTGAGGCTTAGGGCTTACCGGGCCCTTACCATGACAGGTGGCTTATCTTCTACACA 840  
Db 781 CTGAGGCTTAGGGCTTACCGGGCCCTTACCATGACAGGTGGCTTATCTTCTACACA 840  
QY 841 GTCTTACTTCCATGCTGTAACCCCTCTATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
Db 841 GTCTTACTTCCATGCTGTAACCCCTCTATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
QY 901 GCACGTAGGAAGGGGCTGAGCCGCTGACGATTTGGCAGCCAGCAC 945  
Db 901 GCACGTAGGAAGGGGCTGAGCCGCTGACGATTTGGCAGCCAGCAC 945

RESULT 12  
LOCUS AX241793 945 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 541 from Patent W00127158.  
ACCESSION AX241793

VERSION AX241793.1 GI:15798668  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
REFERENCE synthetic construct.  
AUTHORS artificial sequences.  
1 (bases 1 to 945)  
Bellenson, J., Smith, D., Lancet, D., Glusman, G., Fuchs, T. and Yanai, I.  
TITLE Olfactory receptor sequences  
JOURNAL Patent: WO 0127158-A 541 19-APR-2001;  
FEATURES Disinfects (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)  
source location/Qualifiers  
1. 945  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="(H389390 nucleotide)"  
BASE COUNT 175 a 279 c 227 g 264 t  
ORIGIN

Query Match 85.2%; Score 807.4; DB 6; Length 945;  
Best Local Similarity 90.9%; Pred. No. 2.6e-204;  
Matches 859; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 ATGGGAAGATGGGTGAACAGTCTTACACAGATGGCTCTCTCTTGGGCATCTTTTC 60  
Db 1 ATGGAGAGCTGGGTGAACAGTCTTACACAGATGGCTCTCTCTTGGGCATCTTTTC 60  
QY 61 CACAGCCAGACTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
Db 61 CACAGTACTGCTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
QY 121 TGTGGGAATGCT 180  
Db 121 TGTGGGAATGCT 180  
QY 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
Db 181 TACTTCTCTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
QY 241 AAGATGGCAGCCCACT 300  
Db 241 AAGATGGCAGCCCACT 300  
QY 301 CAAATGGCTTTTGTCT 360  
Db 301 CAAATGGCTTTTGTCT 360  
QY 361 TATGACCACTACGTGGCGCTTAGCCACCCTTCACTATCCATCTCTCTCTCTCTCT 420  
Db 361 TATGACCGCTATGTGGCGCTTAGCCACCCTTCACTATCCATCTCTCTCTCTCTCT 420  
QY 421 GTCTGTCTCCAGATTACTGGGAGCTCCCTGGGCTTTGGGATATAGATGAGTGTAG 480  
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCCCTGGGCTTTGGGATATAGATGAGTGTAG 480  
QY 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGGATCACTTTTCTGT 540  
Db 481 ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGGATCACTTTTCTGT 540  
QY 541 GAGGTACAGCTTATGGAAGCTGGGCTGTGACAGACACTTCCCTTTTGAACACCCCTCTC 600  
Db 541 GAGGTACAGCTTATGGAAGCTGGGCTGTGACAGACACTTCCCTTTTGAAGAGGTGATA 600  
QY 601 TTTGCTGTGTCT 660  
Db 601 TTTGCTGTGTCT 660  
QY 661 TGCATCTAGGGGCTGTGCTCCGAATAGCGCTGTGCTGAGGCTTGAAAAAGCCCTGCGC 720  
Db 661 CACATTTCTAGGGGCTGTGCTCCGAATAGCGCTGTGCTGAGGCTTGAAAAAGCCCTGCGC 720  
QY 721 ACCTGCTCTCCACCTTACAGCTGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
Db 721 ACCTGCTCTCCACCTTACAGCTGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780

Db 721 ACCTGCTCCTCCACCTGACAGCTGTCAACCTCTTCTATG6GGGAGCCATGTTCACTAC 780

QY 781 CTGAGCGCTAGCGCTACCGGGCCCTAGCCATGACAGGTGGCCCTTCTTACACA 840

Db 781 CTGAGCGCTAGCGCTACCGGGCCCGCCAGCCATGACAGGTGGCCCTTCTTACACG 840

QY 841 GTCCCTACTCCCATGCTGACCCCTCATTTACAGCTTGAGGAATGGGAGGTATGGGG 900

Db 841 GTCCCTACTCCCATGCTGACCCCTCATTTACAGCTTGAGGAATGGGAGGTATGGGG 900

QY 901 GCACCTGAGGAGGGGCTGACCGCTGACGATTCGACGACGAC 945

Db 901 GCACCTGAGGAGGGGCTGACCGCTGACGATTCGACGACGAC 945

RESULT 13

AC103024 106325 bp DNA linear HTG 13-JUL-2002

LOCUS Rattus norvegicus clone CH230-199E4, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\* 58 unordered pieces.

AC103024

AC103024

VERSION AC103024.3 GI:21730133

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 106325)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 106325)

AUTHORS Worley,K.C.

TITLE JOURNAL

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

Direct Submission

Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 106325)

Worley,K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2002 this sequence version replaced gi:17974430.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GIW

Center clone name: CH230-199E4

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 54989 bases at least Q40

Consensus quality: 61953 bases at least Q30

Consensus quality: 67827 bases at least Q20

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently consists of 58 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1

1014 1013: contig of 1013 bp in length

1114 1113: gap of unknown length

1114 2245: contig of 1132 bp in length

2246 2345: gap of unknown length

2346 3674: contig of 1329 bp in length

3675 3774: gap of unknown length

3775 5318: contig of 1544 bp in length

5319 5418: gap of unknown length

5419 6793: contig of 1375 bp in length

6794 6893: gap of unknown length

6894 8357: contig of 1464 bp in length

8358 8457: gap of unknown length

8458 9678: contig of 1221 bp in length

9679 9778: gap of unknown length

9779 10920: contig of 1142 bp in length

10921 11020: gap of unknown length

11021 12486: contig of 1466 bp in length

12487 12586: gap of unknown length

12587 13730: contig of 1144 bp in length

13731 13830: gap of unknown length

13831 15022: contig of 1192 bp in length

15023 15122: gap of unknown length

15123 16234: contig of 1112 bp in length

16235 16334: gap of unknown length

16335 17872: contig of 1538 bp in length

17873 17972: gap of unknown length

17973 19370: contig of 1398 bp in length

19371 19470: gap of unknown length

19471 20707: contig of 1237 bp in length

20708 20807: gap of unknown length

20808 22976: contig of 2169 bp in length

22977 23077: gap of unknown length

23077 24274: contig of 1198 bp in length

24275 24374: gap of unknown length

24375 25514: contig of 1140 bp in length

25515 25614: gap of unknown length



Query Match 79.18; Score 749.6; DB 2; Length 106325;

*	25615	26694:	contig of 1080 bp in length
*	26695	26794:	gap of unknown length
*	26795	28383:	contig of 1589 bp in length
*	28384	28483:	gap of unknown length
*	28484	29877:	contig of 1394 bp in length
*	29878	29977:	gap of unknown length
*	29978	31516:	contig of 1539 bp in length
*	31517	31616:	gap of unknown length
*	31617	33106:	contig of 1490 bp in length
*	33107	33206:	gap of unknown length
*	33207	34847:	contig of 1641 bp in length
*	34848	34947:	gap of unknown length
*	34948	37050:	contig of 2103 bp in length
*	37051	37150:	gap of unknown length
*	37151	39798:	contig of 2648 bp in length
*	39799	39898:	gap of unknown length
*	39899	41948:	contig of 2050 bp in length
*	41949	42048:	gap of unknown length
*	42049	44144:	contig of 2096 bp in length
*	44145	44244:	gap of unknown length
*	44245	45942:	contig of 1698 bp in length
*	45943	46042:	gap of unknown length
*	46043	47983:	contig of 1941 bp in length
*	47984	48083:	gap of unknown length
*	48084	50236:	contig of 2153 bp in length
*	50237	50336:	gap of unknown length
*	50337	52259:	contig of 1923 bp in length
*	52260	52359:	gap of unknown length
*	52360	53717:	contig of 1358 bp in length
*	53718	53817:	gap of unknown length
*	53818	56007:	contig of 2190 bp in length
*	56008	56107:	gap of unknown length
*	56108	57203:	contig of 1096 bp in length
*	57204	57303:	gap of unknown length
*	57304	59464:	contig of 2161 bp in length
*	59465	59564:	gap of unknown length
*	59565	61655:	contig of 2091 bp in length
*	61656	61755:	gap of unknown length
*	61756	64069:	contig of 2314 bp in length
*	64070	64169:	gap of unknown length
*	64170	65775:	contig of 1606 bp in length
*	65776	65875:	gap of unknown length
*	65876	67680:	contig of 1805 bp in length
*	67681	67780:	gap of unknown length
*	67781	69752:	contig of 1972 bp in length
*	69753	69852:	gap of unknown length
*	69853	71277:	contig of 1425 bp in length
*	71278	71377:	gap of unknown length
*	71378	72946:	contig of 1569 bp in length
*	72947	73046:	gap of unknown length
*	73047	74272:	contig of 1226 bp in length
*	74273	74372:	gap of unknown length
*	74373	76041:	contig of 1669 bp in length
*	76042	76141:	gap of unknown length
*	76142	77614:	contig of 1473 bp in length
*	77615	77714:	gap of unknown length
*	77715	79838:	contig of 2124 bp in length
*	79839	79938:	gap of unknown length
*	79939	82661:	contig of 2723 bp in length
*	82662	82761:	gap of unknown length
*	82762	84760:	contig of 1999 bp in length
*	84761	84860:	gap of unknown length
*	84861	86316:	contig of 1456 bp in length
*	86317	86416:	gap of unknown length
*	86417	88144:	contig of 1728 bp in length
*	88145	88244:	gap of unknown length
*	88245	89772:	contig of 1528 bp in length
*	89773	89872:	gap of unknown length
*	89873	91991:	contig of 2119 bp in length
*	91992	92091:	gap of unknown length
*	92092	94876:	contig of 2785 bp in length

Best Local Similarity 86.9%; Pred. No. 7.1e-189; Matches 824; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY	1	ATGGGAGATGGGTGAACCACTCTACACAGATGGCTTCTCTCTTGGGCACTTTCC	60
Db	56153	ATGGGAATATGGCTGAACGAAATCATCTGTATGATGGGTTCATCTCCTAGGCATCTTCC	56212
QY	61	CACAGCCAGACTGACCTTGCTCTCTCTGCAAGTATAGTGTCTTACAGTGCCCTC	120
Db	56213	CAGAGCCAGACTGACTTCTCTCTCTCTCTGCGGTCAATGTAAGTCTTACAGTGCCCTC	56272
QY	121	TGTGGAAATGCTCTCTCATCTCTCTCATCTACCTGAGCCTGAGCTTCAACCCCATG	180
Db	56273	TGTGGAAATGCTCTCTCATCTCTCTCATCTACCTGATCCCGTCTTCAACCCCATG	56332
QY	181	TACTTCTCTCATGAGCAGCTCTCTCTCATGAGCCTCATGTTGCTGTATAGTGTGCCA	240
Db	56333	TACTTCTCTCATGAGCAGCTCTCTCTCATGAGCCTCATGTTGCTGTATAGTGTGCCA	56392
QY	241	AAGATGGCAGCCACTTCTCTCTCTGAGCAGAGTCCATCTCTTGTGGGCTGTGCCATA	300
Db	56393	AAGATGGCAGTCAACTCTCTCTCTGAGCAGAGTCCATCTCTTGTGGGCTGTGCCATA	56452
QY	301	CAAAATGGCTTTTGTCTCTCTCTGAGGATCTGAGGGCTCTTGTGGCACTCATGGCT	360
Db	56453	CAAAATGGATTTTGTCTCTCTCTGAGGATCTGAGGGCTCTTGTAGCACTCATGGCT	56512
QY	361	TATGACCACATAGTGGCCGTAGCCACCACCTTCACTATCCCATCTCATGAAATCAGAG	420
Db	56513	TATGATCGATATGTGGCCATAGCCACCACCTTCACTATCCCATCTCATGAAATCAGAG	56572
QY	421	GTCTGTCTCCAGATTACTGGGAGCTCTCTGGCCCTTTGGGAAATAGATGAGTGATTCAG	480
Db	56573	GTCTGTCTCCAGATTACTGGGAGTCTCTGGCCCTTTGGGAAATAGATGAGTGATTCAG	56632
QY	481	ATGTGGCAGCCATGGGCTTACTTACTGTGCTCAAGGAGCGTGATCACTTTTCTGT	540
Db	56633	ATGTGGCAGCCATGAGCCTGCTTACTGTGCTCAAGGAGCGTGATCACTTTTCTGT	56692
QY	541	GAGGTACAAGCTTTATTGAAGCTGGCCGTGACAGACATTCCTTTTGGACCCCTCTC	600
Db	56693	GAGGTGCCAGCTTTACTGAAGCTGGCCGTGACAGACATTCCTTTTGGACCCCTCTC	56752
QY	601	TTTGCTTGTGCTTCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	660
Db	56753	TTTGCTTGTGCTTCTTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	56812
QY	661	TGCATCTAGGGGCTGTGCTCCGAATAGCCTGTGCTCAGGCTTGAAAAAGCCCTGGCC	720
Db	56813	CGCATTTTGGGGGCTGTGCTCCGATGACCTGTGCTATTCCTCTCTCTCTCTCTCTCT	56872
QY	721	ACCTGCTCTCCACCTAACAGCTGTCAACCTCTTCTATGAGGGCAGCCATCTCATGTAC	780
Db	56873	ACCTGCTCTCCACCTAACAGCTGTCTCTCTCTATGAGGGCAGCCATCTCATGTAT	56932
QY	781	CTGAGGCTTAGGCGCTACCGGGCCCTAGGCATGACAAAGTGCCCTCTATCTTACACA	840
Db	56933	CTGAGGCGCAAGACGATATCGAGCTCTAGGCATGACAAAGTTGTCTCAATCTTACACA	56992
QY	841	GTCCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG	900
Db	56993	GTCCTTACTTCCATGCTCAACCCCTCATTTATAGCTTGAGAAACAGGAGGTGATGGG	57052
QY	901	GCACTGAGGAAGGGCTGAGCCGCTGACAGATTGGCAGCCAGCACTGA 948	
Db	57053	GCTCTGAGAAAGGCTGGGCCGTTGCAAGGTTGGCAGCCAGCATTTGA 57100	

RESULT 14  
LOCUS AY073214 948 bp DNA linear ROD 03-APR-2002  
DEFINITION Mus musculus olfactory receptor MOR276-1 gene, complete cds.  
ACCESSION AY073214

VERSION	KEYWORDS	GI:18479725
AY073214.1		
SOURCE	Mus musculus.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 948)	
TITLE	Zhang, X. and Firestein, S.	
JOURNAL	The Olfactory receptor gene superfamily of the mouse	
MEDLINE	Nat. Neurosci. 5 (2), 124-133 (2002)	
PUBMED	21676863	
REFERENCE	11802173	
AUTHORS	2 (bases 1 to 948)	
TITLE	Young, J.M., Friedman, C., Williams, E.M., Ross, J.A., Tonnes-Priddy, L. and Trask, B.J.	
JOURNAL	Different evolutionary processes shaped the mouse and human	
MEDLINE	olfactory receptor gene families	
PUBMED	Hum. Mol. Genet. 11 (5), 535-546 (2002)	
REFERENCE	21864068	
AUTHORS	11875048	
TITLE	3 (bases 1 to 948)	
JOURNAL	Adams, M.	
MEDLINE	Direct Submission	
PUBMED	Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,	
REFERENCE	Rockville, MD 20850, USA	
AUTHORS	Location/Qualifiers	
TITLE	1. 948	
JOURNAL	/organism="Mus musculus"	
MEDLINE	/db_xref="taxon:10090"	
PUBMED	/chromosome="11"	
REFERENCE	/map="11B1"	
AUTHORS	/note="GA_x5J8B7W6C8B-402233-401286"	
TITLE	<1. >948	
JOURNAL	/product="Olfactory receptor MOR276-1"	
MEDLINE	1. 948	
PUBMED	/codon_start=1	
REFERENCE	/evidence=not_experimental	
AUTHORS	/product="Olfactory receptor MOR276-1"	
TITLE	/protein_id="AAL60877.1"	
JOURNAL	/db_xref="GI:18479726"	
MEDLINE	/translation="MGIMLNSSVDGFIILGIFSOQTDLLFSTVMLVFTVALCGNV	
PUBMED	LLILILYDPRLHPMYFFLSOLSLMDLVCNIVPKAVNPLSGRKSISFAGCIGOI	
REFERENCE	GFVFSLVGSEGLLGLMAYDRYVALSHPLHYPLMSQKVCLOJAGSSWAFGLDGIIO	
AUTHORS	MVAAMSLPYCGSRIDHFECEVPALLKLACADTSLEFDLLFACVEMLLPFISIIVTS	
TITLE	YARILGAVLRHMSAOSRKALATCSSHLTAVSLFYGAAFIYLRPRRRAPSHDKVVS	
JOURNAL	IFYFTVLPMLNPLIYSLRNREVMGALRKGLDRCRVGSQH"	
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		

QY	301	CAAAATTGGCTTTTGTCTCTCTGTGGGATCGAGGGGCTCTTGCTGGACTCATGGCT	360
Db	301	CAAAATCGGATTTTGTCTCTCTGTGGGATCAGAGGGTCTCTGTAGAGCTATGGCT	360
QY	361	TATGACCACTACGTGGCCGTTAGCCACCACCTTCACTATCCCATCTCATGATTCAGAGG	420
Db	361	TATGATCGCTAATGEGCCATTAGCCACCACCTTCACTATCCCATCTCATGAGCCAAAG	420
QY	421	GTCGTCTCCAGATTACTGGGAGCTCTGGGCCCTTTGGGATAATAGATGAGTGAATTGAG	480
Db	421	GTCGTCTCCAGATTCTGGAAGTCTGGGCTTTGGGATCCTTGATGAATAATTGAG	480
QY	481	ATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGGCGTGGATCACTTTTCTGT	540
Db	481	ATGGTGGCAGCCATGAGCCTGCCCTACTGTGGCTCAAGGTATATGATCACTTCTCTGT	540
QY	541	GAGGTACAAGCTTTATGAAAGCTGGCCGTGCAGACACTTCCCTTTTGACACCCTCTC	600
Db	541	GAGGTGGCCGCTTACTGAAGCTGGCCTGTGCAGACACTCCCTTTCGACACCCTGCTC	600
QY	601	TTTGCTTGTCTGTCTTCATGCTTCTTCCCTTCCATCATCATGAGCCCTCTATGCT	660
Db	601	TTTGTCTGTGTCTTATGCTGTCTTCTTCCATCATCATGAGCTTCTCTATGCT	660
QY	661	TGCATCTAGGGGCTGTGCTCCGAATAGCTCTGTCTCAGGCCCTGAAAAAAGCCCTGGCC	720
Db	661	CGCATCTGGGGGCTGTGCTCCGATAGCACTCTGCCAGTCCCGAAAAAAGGCCCTGGCC	720
QY	721	ACCTGCTCTCCACCACTAACAGCTGTCAACCCTCTTATAGGGGACAGCATGTTCACTGAC	780
Db	721	ACTTGTCTCTCCACCACTGACAGCTGTCTCTCTCTTCTACGGGGACAGCAATGTTCACTGAC	780
QY	781	CTGAGGCGCTAGGCGCTACCGGGGCCCTAGACCATGACAAGGTGGCTCTATCTTACACA	840
Db	781	CTGAGGCCAACGGCATAGCGCGCTCTAGACCATGACAAAGTGTCTCAATCTTCTACACA	840
QY	841	GTCCTTACTCCATGCTGAACCCCTCAATTTACAGCTTGAGGAATGGGAGGTGATGGGG	900
Db	841	GTTCTTACTCTATGCTCAACCCCTCAATTTATAGCTTGAGGAACAGGAGGTGATGGGG	900
QY	901	GCACGTAGAGGAGGGCTGACCCGCTGACGATTTGGCAGCCAGCACTGA 948	
Db	901	GCACGTAGAGAAAGACTGACCGCTGACAGGGTTGGCAGCCAAACATTTGA 948	

RESULT 15	
AL645688	
LOCUS	123115 bp
DEFINITION	Mouse DNA sequence from clone RP23-54N20 on chromosome 11, complete
	AL645688
	DNA linear
	ROD 05-APR-2002

ACCESSION	sequence.
VERSION	AL645688
KEYWORDS	AL645688.12
SOURCE	HTG.
ORGANISM	house mouse.
	Mus musculus

REFERENCE	1
AUTHORS	Howden, P.
TITLE	Direct Submission
JOURNAL	Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,

COMMENT

On Mar 20, 2002 this sequence is from a different clone. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 20:23:29 ; Search time 267 Seconds  
(without alignments)  
7995.859 Million cell updates/sec

Title: US-09-634-109D-1  
Perfect score: 948  
Sequence: 1 atggggaagatggtgaacca.....ggattgcagccagcactga 948

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	948	100.0	948	22	AAS00555 Human CON167 G pro
2	948	100.0	1070	22	AAS00560 Human CON167 G pro
3	940	99.2	951	24	ABK97214 Human G-protein co
4	935.4	98.7	1120	24	ABQ88046 Human GPCRX cDNA #
5	908.8	95.9	975	22	AAH31969 Human olfactory re
6	810.4	85.5	948	22	AAS42296 Human cDNA encodin
7	810.4	85.5	948	24	ABK68506 Human DNA for olfa
8	810.4	85.5	948	24	AAD29683 Human G-protein co
9	810.4	85.5	948	24	ABK37582 DNA encoding G-con

10	809	85.3	945	22	AAH31875 Human olfactory re
11	807.4	85.2	945	22	AAH31968 Human olfactory re
12	672.4	70.9	965	24	AAD26374 Human G-protein co
13	526.6	55.5	669	21	AAH51179 Murine olfactory r
14	526.6	55.5	669	21	AAH51183 Murine olfactory r
15	525	55.4	669	21	AAH51181 Murine olfactory r
16	390.2	41.2	488	22	AAH84008 Gorilla olfactory
17	389.4	41.1	939	24	ABK68668 Human DNA for olfa
18	388.8	41.0	1290	19	AAV53208 Human olfactory OL
19	386.2	40.7	936	24	ABK68667 Human DNA for olfa
20	379.8	40.1	939	24	ABK68670 Human DNA for olfa
21	371	39.1	963	24	ABK68671 Human DNA for olfa
22	366.2	38.6	963	24	ABK68669 Human DNA for olfa
23	362.2	38.2	984	22	ABN89126 Human GPCR7c nucle
24	361.8	38.2	975	22	AAS42423 Human cDNA encodin
25	361.8	38.2	975	24	ABK68567 Human DNA for olfa
26	361.8	38.2	975	24	AAD29682 Human G-protein co
27	361.8	38.2	975	24	ABK37709 Human cDNA encodin
28	360.6	38.0	951	24	AAS42447 Human GPCR7a nucle
29	360.6	38.0	951	24	ABN89124 Human GPCR7a nucle
30	360.6	38.0	951	24	ABK37733 Human olfactory re
31	352.4	37.2	964	22	AAH31643 Human GPCR7b nucle
32	351.8	37.1	992	24	ABN89125 Human DNA for olfa
33	349.6	36.9	870	24	ABK68556 Human GPCR6a nucle
34	346.6	36.6	948	24	ABN89121 Human olfactory re
35	345	36.4	945	22	AAH31912 Human cDNA encodin
36	345	36.4	948	22	AAS42426 Human GPCR6c nucle
37	345	36.4	948	24	ABN89123 Human DNA for olfa
38	345	36.4	948	24	ABK68568 Human DNA for olfa
39	345	36.4	948	24	ABK37712 Human encoding G-con
40	345	36.4	949	24	ABN89122 Human GPCR6b nucle
41	344.8	36.4	963	24	ABO88381 Human G-protein co
42	343.8	36.3	990	22	AAH31892 Human olfactory re
43	343.8	36.3	993	24	ABK68569 Human DNA for olfa
44	343.8	36.3	1047	22	AAS42425 Human cDNA encodin
45	343.8	36.3	1047	24	ABK37711 DNA encoding G-con

ALIGNMENTS

RESULT 1					
AAS00555	ID	AAS00555	standard;	DNA; 948 BP.	
XX	AC	AAS00555;			
XX	DT	14-MAY-2001	(first entry)		
XX	DE	Human CON167 G protein seven transmembrane receptor DNA #1.			
KW	CON167;	G protein coupled receptor; transmembrane receptor; ds;			
KW	neurological disorder; psychiatric disease; schizophrenia; depression;				
KW	anxiety; bipolar disorder; affective disorder; epilepsy; neuritis;				
KW	attention deficit hyperactivity disorder; ADHD; neurasthenia; neuropathy;				
KW	neurosis; Alzheimer's disease; Parkinson's disease; migraine; therapy;				
KW	senile dementia; hybridisation assay; diagnosis; transgenic animal.				
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	1..948				
FT	CDS	/*tag= a			
FT		/product= "CON167"			
XX					
PN	WO200114554-A1.				
XX					
PD	01-MAR-2001.				
XX					
PF	08-AUG-2000; 2000WO-US21566.				
XX					
PR	19-AUG-1999; 99US-0377563.				
XX					

PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Vogeli G, Wood LS;  
 XX  
 DR WPI: 2001-218450/22.  
 DR P-PSDB; AAU00529.  
 XX  
 PT Novel purified isolated seven transmembrane receptor polypeptide  
 PT (G-protein coupled receptor) useful for treating neurological and  
 PT psychiatric diseases such as schizophrenia, depression, anxiety,  
 PT bipolar disease and affective disorder  
 XX  
 PS Claim 6; Page 62-63; 72pp; English.  
 CC The sequence represents DNA encoding human CON167 G protein seven  
 CC transmembrane receptor. The DNA and the protein are useful for modulating  
 CC activity of CON167 in a mammal comprising cells that express CON167,  
 CC preferably in a human suffering from a neurological disorder and/or  
 CC psychiatric diseases such as schizophrenia, depression, anxiety, bipolar  
 CC disease, affective disorder, attention deficit hyperactivity disorder  
 CC (ADHD), epilepsy, neuritis, neuroasthenia, neuropathy, neuroses,  
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
 CC The polynucleotides are useful in hybridisation assays to detect the  
 CC capacity of cells to express CON167, for large scale expression of CON167  
 CC required for research, for identification and isolation of new  
 CC polynucleotides encoding the related CON167 polypeptides, in diagnostic  
 CC methods for identifying any genetic alteration in a CON167 locus that  
 CC underlies a disease state (which is useful for selection of therapeutic  
 CC strategies), and for the development of transgenic animals that fail to  
 CC express functional CON167 or that express a variant of CON167.  
 XX  
 SQ Sequence 948 BP; 172 A; 278 C; 230 G; 268 T; 0 other;  
 Query Match 100.0%; Score 948; DB 22; Length 948;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-256;  
 Matches 948; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GAGGTACAGCTTTATGAGCTGGCGCTGACAGACACTTCCCTTTTGACACCCCTCCTC 600  
 Db 541 GAGGTACAGCTTTATGAGCTGGCGCTGACAGACACTTCCCTTTTGACACCCCTCCTC 600  
 QY 601 TTTGCTTCTGTGTCTCATGCTTCTCCCTCCCTTCCATCATGAGCCCTCATGCT 660  
 Db 601 TTTGCTTCTGTGTCTCATGCTTCTCCCTCCCTTCCATCATGAGCCCTCATGCT 660  
 QY 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTCTGTACAGCCCTGAAAGGCTGCGC 720  
 Db 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTCTGTACAGCCCTGAAAGGCTGCGC 720  
 QY 721 ACCTGCTCCCTCCACCTAACAGCTGTACCCCTCTTATGGGACAGCATGTTATGATAC 780  
 Db 721 ACCTGCTCCCTCCACCTAACAGCTGTACCCCTCTTATGGGACAGCATGTTATGATAC 780  
 QY 781 CTGAGGCTAGGGGCTACCGGGCCCTAGCATGACAGGCTGCGCTTATCTTACACA 840  
 Db 781 CTGAGGCTAGGGGCTACCGGGCCCTAGCATGACAGGCTGCGCTTATCTTACACA 840  
 QY 841 GTCTTACTCCCATGCTGACACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
 Db 841 GTCTTACTCCCATGCTGACACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
 QY 901 GCACTGAGAGAGGGGCTGACCGCTGACAGATTGGACCCAGCACTGA 948  
 Db 901 GCACTGAGAGAGGGGCTGACCGCTGACAGATTGGACCCAGCACTGA 948

RESULT 2  
 AAS00560  
 ID AAS00560 standard; DNA: 1070 BP.  
 XX  
 AC AAS00560;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX  
 DE Human CON167 G protein seven transmembrane receptor DNA #2.  
 XX  
 KW CON167; G protein coupled receptor; transmembrane receptor; ss;  
 KW neurological disorder; psychiatric disease; schizophrenia; depression;  
 KW anxiety; bipolar disorder; affective disorder; epilepsy; neuritis;  
 KW attention deficit hyperactivity disorder; ADHD; neuroasthenia; neuropathy;  
 KW neurosis; Alzheimer's disease; Parkinson's disease; migraine; therapy;  
 KW senile dementia; hybridisation assay; diagnosis; transgenic animal;  
 KW 5' untranslated region.  
 KW  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..122  
 FT CDS /\*tag= a  
 FT 123..1070  
 FT /\*tag= b  
 FT /\*product= "CON167"  
 PN WO200114554-A1.  
 XX  
 PD 01-MAR-2001.  
 XX  
 PF 08-AUG-2000; 2000WO-US21566.  
 XX  
 PR 19-AUG-1999; 99US-0377563.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Vogeli G, Wood LS;  
 XX  
 DR WPI: 2001-218450/22.  
 XX  
 PT Novel purified isolated seven transmembrane receptor polypeptide  
 PT (G-protein coupled receptor) useful for treating neurological and  
 PT psychiatric diseases such as schizophrenia, depression, anxiety,



PT bipolar disease and affective disorder -  
XX  
PS Example 6; Page 67-68; 72pp; English.  
XX  
CC The sequence represents DNA including the original 5' untranslated  
CC region, encoding human CON167 G protein seven transmembrane receptor.  
CC The DNA and the protein are useful for modulating activity of CON167 in  
CC a mammal comprising cells that express CON167, preferably in a human  
CC suffering from a neurological disorder or psychiatric diseases such as  
CC schizophrenia, depression, anxiety, bipolar disease, affective disorder,  
CC attention deficit hyperactivity disorder (ADHD), epilepsy, neuritis,  
CC neurasthenia, neuropathy, neuroses, Alzheimer's disease, Parkinson's  
CC disease, mygraine and senile dementia. The polynucleotides are useful in  
CC hybridisation assays to detect the capacity of cells to express CON167,  
CC for large scale expression of CON167 required for research, for  
CC identification and isolation of new polynucleotides encoding the related  
CC CON167 polypeptides, in diagnostic methods for identifying any genetic  
CC alteration in a CON167 locus that underlies a disease state (which is  
CC useful for selection of therapeutic strategies), and for the development  
CC of transgenic animals that fail to express functional CON167 or that  
CC express a variant of CON167.  
XX  
SQ Sequence 1070 BP; 207 A; 305 C; 254 G; 304 T; 0 other;

Query Match 100.0%; Score 948; DB 22; Length 1070;  
Best Local Similarity 100.0%; Pred. No. 5.5e-256;  
Matches 948; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGATGGGTGAACCAAGTCTACACAGATGGCTCTCTCTGGGCATCTTTCC 60  
Db 123 ATGGAGATGGGTGAACCAAGTCTCTACACAGATGGCTCTCTCTGGGCATCTTTCC 182  
QY 61 CACAGCCAGACTGACCTTGCTCTCTCTCTGAGTATAGTGTGTCTTCACAGTGGCCCTC 120  
Db 183 CACAGCCAGACTGACCTTGCTCTCTCTCTGAGTATAGTGTGTCTTCACAGTGGCCCTC 242  
QY 121 TGTGGAAATGCTCTCTCATCTTCTCATCTTACCTGAGCGCTGGACTTCACACCCCATG 180  
Db 243 TGTGGAAATGCTCTCTCATCTTCTCATCTTACCTGAGCGCTGGACTTCACACCCCATG 302  
QY 181 TACTTCTCTCAGCCAGCTCTCCCTCATAGGACCTCATAGTTGTCTTAACATTTGCCA 240  
Db 303 TACTTCTCTCAGCCAGCTCTCCCTCATAGGACCTCATAGTTGTCTTAACATTTGCCA 362  
QY 241 AAGATGGCAGCCAACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATA 300  
Db 363 AAGATGGCAGCCAACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATA 422  
QY 301 CAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTGTGGGACTCATGGCT 360  
Db 423 CAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTGTGGGACTCATGGCT 482  
QY 361 TATGACCACTACGTGGCCGTAGCCACCCACTTCACATATCCCATCTCATGAATCAGAGG 420  
Db 483 TATGACCACTACGTGGCCGTAGCCACCCACTTCACATATCCCATCTCATGAATCAGAGG 542  
QY 421 GTCTGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATAATAGATGAGTGATTGAG 480  
Db 543 GTCTGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATAATAGATGAGTGATTGAG 602  
QY 481 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGAGCGTGATCACTTTTCTGT 540  
Db 603 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGAGCGTGATCACTTTTCTGT 662  
QY 541 GAGGTACAAGCTTTATGAAGCTGGCCGTGCAGACACTTCCCTTTTGAACACCCTCTC 600  
Db 663 GAGGTACAAGCTTTATGAAGCTGGCCGTGCAGACACTTCCCTTTTGAACACCCTCTC 722  
QY 601 TTTGCTGTGTGTCTTCATGTCTTCTCTTCCCTTCTCCATCATCATGGCCCTCTATGCT 660  
Db 723 TTTGCTGTGTGTCTTCATGTCTTCTCTTCCCTTCTCCATCATCATGGCCCTCTATGCT 782  
QY 661 TGCATCTTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCCCTGGAAGAAAGCCCTG 720

Db 783 TGCATCTTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCCCTGGAAGAAAGCCCTG 842  
QY 721 ACCTGTCTCTCCACACCTAACAAGCTGTACACCCCTCTTATATGGGGCAGCCATGTTATG 780  
Db 843 ACCTGTCTCTCCACACCTAACAAGCTGTACACCCCTCTTATATGGGGCAGCCATGTTATG 902  
QY 781 CTGAGGCCTTAGGGCGCTACCGGGCCCTAGCCATGACAGAGGTGGCCCTTATCTTACACA 840  
Db 903 CTGAGGCCTTAGGGCGCTACCGGGCCCTAGCCATGACAGAGGTGGCCCTTATCTTACACA 962  
QY 841 GTCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAAGGATGGGAGGTGATGGG 900  
Db 963 GTCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAAGGATGGGAGGTGATGGG 1022  
QY 901 GCACGTGAGGAAGGGGCTTGACCCGCTGACGATTGGCAGCCAGCACTGA 948  
Db 1023 GCACGTGAGGAAGGGGCTTGACCCGCTGACGATTGGCAGCCAGCACTGA 1070

RESULT 3

ABK97214  
ID ABK97214 standard; cDNA; 951 BP.  
XX  
AC ABK97214;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GPCR) GPCR gene #8.  
XX  
KW G-protein coupled receptor; receptor; GPCR; GPCR; cardiomyopathy;  
KW atherosclerosis; diabetes; cell signal processing; cancer; trauma;  
KW metabolic pathway modulation; neuro-olfactory system; surgery;  
KW neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;  
KW uterus cancer; immune response; acquired immunodeficiency syndrome;  
KW AIDS; asthma; Crohn's disease; multiple sclerosis;  
KW Albright hereditary osteodystrophy; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200250117-A2.  
XX  
PD 27-JUN-2002.  
XX  
PF 18-DEC-2001; 2001WO-US49077.  
XX  
PR 18-DEC-2000; 2000US-256635P.  
PR 21-DEC-2000; 2000US-257876P.  
PR 04-JAN-2001; 2001US-259743P.  
PR 10-JAN-2001; 2001US-260718P.  
PR 12-JAN-2001; 2001US-261498P.  
PR 24-JAN-2001; 2001US-263689P.  
PR 08-FEB-2001; 2001US-267464P.  
PR 22-FEB-2001; 2001US-271021P.  
PR 14-MAR-2001; 2001US-275946P.  
PR 23-MAR-2001; 2001US-278150P.  
PR 18-APR-2001; 2001US-284591P.  
PR 23-APR-2001; 2001US-285718P.  
PR 19-JUN-2001; 2001US-299327P.  
PR 16-AUG-2001; 2001US-312902P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;  
PI Vernet CAM, Li L, Shenoy S, Casman SJ;  
XX  
DR WPI; 2002-528447/56.  
DR P-PSDB; ABG68141.  
XX  
PT New G-protein coupled receptor polypeptides for treating or preventing  
PT cardiomyopathy, atherosclerosis, diabetes, multiple sclerosis, acquired  
PT immunodeficiency syndrome or cancer in humans



PS Claim 5; Page 86; 110pp; English.  
XX  
CC The present invention relates to a new G-protein coupled receptor (GPCRX)  
CC polypeptide. The GPCR polypeptide, GPCR nucleic acid and antibody are  
CC useful for treating, preventing or alleviating a GPCR-associated  
CC disorder or a pathological state in a subject, particularly a human. In  
CC particular, the disorder is cardiomyopathy, atherosclerosis, diabetes, or  
CC a disorder related to cell signal processing and metabolic pathway  
CC modulation. The GPCR polypeptide and nucleic acid are also useful for  
CC diagnosing the presence of or predisposition to a disease associated  
CC with altered levels of GPCR, particularly cancer. The GPCR nucleic acid  
CC and polypeptide are especially useful in therapeutic or prophylactic  
CC applications for disorders of the neuro-olfactory system, e.g. those  
CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding  
CC the protein is useful in gene therapy for treating the above conditions.  
CC Furthermore, the nucleic acids and polypeptides are useful in treating  
CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune  
CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's  
CC disease, multiple sclerosis or Albiglight hereditary osteodystrophy. These  
CC are also useful in developing powerful assay system for functional  
CC analysis of various human disorders, as well as in diagnostic  
CC applications. The present nucleic acid sequence represents one of a  
CC collection (ABK97207-ABK97226) of human GPCR genes that encode the  
CC human GPCR proteins (AAU68134-AAU68153) of the invention.  
CC  
XX  
SQ Sequence 951 BP; 172 A; 278 C; 231 G; 270 T; 0 other;

Query Match 99.28; Score 940; DB 24; Length 951;  
Best Local Similarity 99.58; Pred. No. 9.2e-254;  
Matches 943; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAGATGGGTGAACAGTCTACACAGATGGCTTCTTCTTGGGATCTTTTC 60  
Db 2 ATGGAGATGGGTGAACAGTCTACACAGATGGCTTCTTCTTGGGATCTTTTC 61  
QY 61 CACAGCCAGACTGACCTTGTCTCTCTGAGTATGATGCTTACAGTGGCCCTC 120  
Db 62 CACAGCCAGACTGACCTTGTCTCTCTGAGTATGATGCTTACAGTGGCCCTC 121  
QY 121 TGTGGATGCTCTCTCATCTCTCATCTGAGGCTGAGCTTACACCCCATG 180  
Db 122 TGTGGATGCTCTCTCATCTCTCATCTGAGGCTGAGCTTACACCCCATG 181  
QY 181 TACTTCTCTCAGCCAGCTCTCCCTCATGAGCTCATGTGCTGTAACTTTGCA 240  
Db 182 TACTTCTCTCAGCCAGCTCTCCCTCATGAGCTCATGTGCTGTAACTTTGCA 241  
QY 241 AAGATGGCAGCCAACTCTCTGTGGAGAGTCCATCTCTTGTGGGCTGTGCATA 300  
Db 242 AAGATGGCAGCCAACTCTCTGTGGAGAGTCCATCTCTTGTGGGCTGTGCATA 301  
QY 301 CAATTTGGCTTTTGTCTCTGTGGAGTGAAGGCTCTGTGGGCTCATGGCT 360  
Db 302 CAATTTGGCTTTTGTCTCTGTGGAGTGAAGGCTCTGTGGGCTCATGGCT 361  
QY 361 TATGACCACTACGTGGCCGTAGCCACCCACTTCACTATCCATCCTCATGATCAGAG 420  
Db 362 TATGACCACTACGTGGCCGTAGCCACCCACTTCACTATCCATCCTCATGATCAGAG 421  
QY 421 GTCTGTCTCAGATTACTGGAGCTCTGGGCTTTGGGATATAGATGAGTGAATTGAG 480  
Db 422 GTCTGTCTCAGATTACTGGAGCTCTGGGCTTTGGGATATAGATGAGTGAATTGAG 481  
QY 481 ATGGTGGCAGCCATGGGCTTACTGTGCTCAAGAGCGTGATCATTCTTCTGT 540  
Db 482 ATGGTGGCAGCCATGGGCTTACTGTGCTTGAAGAGCGTGATCATTCTTCTGT 541  
QY 541 GAGGTACAAAGCTTTATGAAGCTGGCCGTGAGACACTTCCCTTTTGAACACCTCCTC 600  
Db 542 GAGGTACAAAGCTTTATGAAGCTGGCCGTGAGACACTTCCCTTTTGAACACCTCCTC 601  
QY 601 TTTGCTTGTGTCTTCAAGCTTCTCTTCCCTTCTCATCATCATGAGGCTCTATGCT 660  
XX

Db 602 TTGCTTGCTGTGTCTTCATGCTTCTTCTTCCCTTCCCATCATGAGCCCTCATGCT 661  
QY 661 TGCATCTTAGGGGCTGTGCTCCGAATACGCTTGTCTCAGGCGCTGAAAAAGCCCTGCCC 720  
Db 662 TGCATCTTAGGGGCTGTGCTCCGAATACGCTTGTCTCAGGCGCTGAAAAAGCCCTGCCC 721  
QY 721 ACCTGCTCTCCACCTAACAGCTGTACACCTCTTCTATGCGGACGCCATGTTCATGTAC 780  
Db 722 ACCTGCTCTCCACCTAACAGCTGTACACCTCTTCTATGCGGACGCCATGTTCATGTAC 781  
QY 781 CTGAGGCGCTAGGCGCTACCGGCGCCCTAGCCATGACACAGTGGCCTTACTTCTACACA 840  
Db 782 CTGAGGCGCTAGGCGCTACCGGCGCCCTAGCCATGACACAGTGGCCTTACTTCTACACA 841  
QY 841 GTCCTTACTCCATGTCTGAACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 900  
Db 842 GTCCTTACTCCATGTCTGAACCCCTCATTTACAGCTTGAAGATGGGAGGTGATGGG 901  
QY 901 GCACTGAGGAGGGGCTGACCGCTGCAGGATTTGGCAGCCAGCACTGA 948  
Db 902 GCACTGAGGAGGGGCTGACCGCTGCAGGATTTGGCAGCCAGCACTGA 949

RESULT 4  
ABQ88046  
ID ABQ88046 standard; cDNA; 1120 BP.  
XX AC ABQ88046;  
XX AC 18-SEP-2002 (first entry)  
DT 18-SEP-2002 (first entry)  
XX DE Human GPCR cDNA #4.  
XX DE Human; G-protein coupled receptor; GPCR; cardiac; antiatherosclerotic;  
KW antidiabetic; hypotensive; anorectic; cytosstatic; immunosuppressive;  
KW neuroprotective; nootropic; antiparkinsonian; antilipemic; gene therapy;  
KW cardiomyopathy; atherosclerosis; cell signal processing; diabetes;  
KW metabolic pathway; hypertension; obesity; cancer; Alzheimer's disease;  
KW graft versus host disease; Parkinson's disorder; dyslipidaemia; gene; ss.  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 3..794 /\*tag= a  
FT CDS /product= "GPCRX"  
FT  
XX WO200255557-A2.  
XX PD 18-JUL-2002.  
XX PF 18-DEC-2001; 2001WO-US49111.  
XX PR 18-DEC-2000; 2000US-256635P.  
XX PR 21-DEC-2000; 2000US-257876P.  
XX PR 04-JAN-2001; 2001US-259743P.  
XX PR 10-JAN-2001; 2001US-260718P.  
XX PR 12-JAN-2001; 2001US-261498P.  
XX PR 24-JAN-2001; 2001US-263689P.  
XX PR 08-FEB-2001; 2001US-267464P.  
XX PR 22-FEB-2001; 2001US-271021P.  
XX PR 14-MAR-2001; 2001US-275946P.  
XX PR 23-MAR-2001; 2001US-278150P.  
XX PR 16-AUG-2001; 2001US-312902P.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;  
XX PI Vernet CAM, Li L, Shenoy S, Casman SJ, Gusev V;  
XX WP1; 2002-500839/53.  
XX DR P-PSDB; ABP61132.  
XX





QY 181 TACTTCTCTCAGCCAGCTCTCCCTCATGAGCTCATGTTGCTGTACATTGTGCCA 240  
 |||||||  
 Db 181 TACTTCTCTCAGCCAGCTCTCCCTCATGAGCTCATGTTGCTGTACCAATGTGCCA 240  
 QY 241 AAGATGGCAGCCAACTTCTCTGCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGCCATA 300  
 |||||||  
 Db 241 AAGATGGCAGCCAACTTCTCTGCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGCCATA 300  
 QY 301 CAAATTGGCTTTTGTCTCTCTGTGGATGAGGGGCTTGTGCTGGACTCATGGCT 360  
 |||||||  
 Db 301 CAAATTGGCTTTTGTCTCTCTGTGGATGAGGGGCTTGTGCTGGACTCATGGCT 360  
 QY 361 TATGACCACTAGCTGGCCGTTAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 420  
 |||||||  
 Db 361 TATGACCACTAGCTGGCCATGAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 420  
 QY 421 GTCGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGATATAGATGAGTGAATTCAG 480  
 |||||||  
 Db 421 GTCGTCTCCAGATTACTGGGAGCTCTGGGCTTTGGATATAGATGAGTGAATTCAG 480  
 QY 481 ATGGTGGCAGCATGGGCTTACCTTACTGTGCTCAAGAGCGTGGATCATTCTTCTGT 540  
 |||||||  
 Db 481 ATGGTGGTATGATTAATTTCCCTTACTGTGCTTGAAGAGTGAACCATTTCTTCTGT 540  
 QY 541 GAGGTACAAAGCTTTATGAAAGCTGGGCTGTGAGACATCCCTTTTGGACACCTCTC 600  
 |||||  
 Db 541 GAGATGCTATCTGTGAAAGCTGGGCTGTGAGACATCCCTTTTGGAGAGGTGATA 600  
 QY 601 TTTGCTTGTCTGTCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
 |||||||  
 Db 601 TTTGCTTGTCTGTCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
 QY 661 TGCATCTAGGGGCTGTGCTCCGATATACGCTGCTCAGGCTGGAAGGCTGGGCC 720  
 |||||  
 Db 661 CACATTTAGGAGCTGTGCTGCAATGACATCTGCTCAGGCTGGAAGGCTGGGCC 720  
 QY 721 ACCTGCTCTCCACCTTAACAGCTGTACACCTCTCTCTCTCTCTCTCTCTCTCTCTAC 780  
 |||||||  
 Db 721 ACCTGCTCTCCACCTGTACAGCTGTACACCTCTCTCTCTCTCTCTCTCTCTCTAC 780  
 QY 781 CTGAGGCTAGGGGCTACCGGCTCTAGCCATGACAAGGTGGCTCTATCTTCTACACA 840  
 |||||||  
 Db 781 CTGAGGCTAGGGGCTACCGGCTCTAGCCATGACAAGGTGGCTCTATCTTCTACACG 840  
 QY 841 GTCCTTACTCCATGCTGACCCCTCATTTTACAGCTTGAAGAAATGGGAGGTGATGGG 900  
 |||||||  
 Db 841 GTCCTTACTCCATGCTGACCCCTCATTTTACAGCTTGAAGAAATGGGAGGTGATGGG 900  
 QY 901 GCACTGAGGAAGGGGCTGAGCGCTGACAGATGAGGAGCCAGCACTGA 948  
 |||||||  
 Db 901 GCACTGAGGAAGGGGCTGAGCGCTGACAGATGAGGAGCCAGCACTGA 948  
 RESULT 7  
 ABK68506  
 ID ABK68506 standard; DNA; 948 BP.  
 XX  
 AC ABK68506;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Human DNA for olfactory and pheromone G protein-coupled receptor #106.  
 XX  
 KW Human; ds; gene; olfactory and pheromone G protein coupled receptor;  
 KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;  
 KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
 KW sterility; psychotic disorder; neurological disorder; anxiety;  
 KW schizophrenia; manic depression; depression; axonal growth;  
 KW menstrual cycle; appetite sexual motivation; sexual attraction;  
 KW aggression.  
 XX  
 OS Homo sapiens.

XX  
 PN WO200224726-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 21-SEP-2001; 2001WO-BE00162.  
 XX  
 PR 22-SEP-2000; 2000EP-0870211.  
 XX  
 PA (CHEM-) CHEMCOM SA.  
 XX  
 PI Veithen A;  
 XX  
 DR WPI; 2002-330013/36.  
 DR P-PSDB; AAU95619.  
 XX  
 PT Novel pheromone G-protein coupled receptor and receptor-derived  
 PT agonists, antagonists or inhibitors useful in food or cosmetic products  
 PT or in the treatment or prevention of neurological disorders such as  
 PT anxiety and schizophrenia.  
 XX  
 PS Disclosure; Page 330-331; 833pp; English.  
 XX  
 CC The invention relates to olfactory and pheromone G-protein coupled  
 CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
 CC portion and its encoding polynucleotide. Also included are an agonist,  
 CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector  
 CC comprising the polynucleotide, a cell transformed by the vector, a non-  
 CC human mammal comprising a partial or total deletion of the polynucleotide  
 CC encoding the receptor and screening (detection and possibly, recovering)  
 CC of compounds which are known or not known to be agonist, antagonists or  
 CC inhibitors of natural compounds to the GPCR. The receptor-derived  
 CC agonists, antagonists, inhibitors or compounds are used as an  
 CC improvement, elimination or substitution of an existing taste and/or a  
 CC fragrance of (or in) the food and/or cosmetic products. They can also be  
 CC used in the preparation of medicament in the treatment and/or prevention  
 CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
 CC neurological disorders, including anxiety, schizophrenia, manic  
 CC depression, depression, for promoting axonal growth, nerve cell  
 CC connection and nerve regeneration for modulating male and female  
 CC endocrine functions, hormone production and the menstrual cycle, for the  
 CC prevention or the treatment by stimulation of several mammalian  
 CC behaviours, such as stimulation or suppression of appetite, sexual  
 CC motivation, sexual attraction, aggression and for promoting or  
 CC suppressing chemical communication between organisms. The present  
 CC sequence is a human DNA encoding an olfactory and pheromone GPCR.  
 XX  
 SQ Sequence 948 BP; 176 A; 279 C; 228 G; 265 T; 0 other;  
 Query Match 85.5%; Score 810.4; DB 24; Length 948;  
 Best Local Similarity 90.9%; Pred. No. 2.3e-217;  
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
 QY 1 ATGGGAAGATGGGTGAACCAAGCTTACACAGATGGCTTCTCTTGGCATCTTTCC 60  
 |||||  
 Db 1 ATGGAGACGTGGGTGAACCAAGCTTACACAGATGGCTTCTCTTAGGCATCTTCTCC 60  
 QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGAGTTATGTGTCTTACAGTGGCCCTC 120  
 |||||  
 Db 61 CACAGTACTGTGACCTTGTCTCTCTCTCTGAGTTATGTGTCTTACAGTGGCCCTC 120  
 QY 121 TGTGGGAATGCTCTCTCATCTCTCTCATCTGAGCGTGGACTTACACCCCATG 180  
 |||||||  
 Db 121 TGTGGGAATGCTCTCTCATCTCTCTCATCTGAGCCCTCACACCCCATG 180  
 QY 181 TACTTCTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTGTGTGTAACTGTGCA 240  
 |||||||  
 Db 181 TACTTCTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTGTGTGTAACTGTGCA 240  
 QY 241 AAGATGGCAGCCAACTTCTCTGCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGCCATA 300  
 |||||||  
 Db 241 AAGATGGCAGCCAACTTCTCTGCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGCCATA 300





Db 61 CACAGTACTGCTGACCTGTGCTCTCTCTCCGCTGATGAGGCTCTTACACAGTGGCCCTC 120  
 QY 121 TGTGGGAATGTCCTCTCATCTTCTCATCTTACCTGAGCGCTGACCTTACACCCCATG 180  
 Db 121 TGTGGGAATGTCCTCTCATCTTCTCATCTTACCTGAGCGCTTACACCCCATG 180  
 QY 181 TACTTCTTCTCAGCCAGCTCTCTCATCTTACCTGAGCGCTTACCTGAGCGCTTAC 240  
 Db 181 TACTTCTTCTCAGCCAGCTCTCTCATCTTACCTGAGCGCTTACCTGAGCGCTTAC 240  
 QY 241 AAGATGGCAGCACTCTCTCTCTGAGGAGTCCATCTCTTGTGGGCTGTGGCATA 300  
 Db 241 AAGATGGCAGCACTCTCTCTCTGAGGAGTCCATCTCTTGTGGGCTGTGGCATA 300  
 QY 301 CAAATTTGGCTTTTGTCTCTCTGAGGAGTCCATCTCTTGTGGGCTGTGGCATA 360  
 Db 301 CAAATTTGGCTTTTGTCTCTCTGAGGAGTCCATCTCTTGTGGGCTGTGGCATA 360  
 QY 361 TATGACCACTAGCTGAGCGCTTACCCACCTCATCTTACCTCATGAAATCAGAGG 420  
 Db 361 TATGACCGCTATGCTGAGCGCTTACCCACCTCATCTTACCTCATGAAATCAGAGG 420  
 QY 421 GTCCTGCTCCAGATTTACTGAGGAGCTCTGAGGAGTCCATCTTGTGGAGATGATTCAG 480  
 Db 421 GTCCTGCTCCAGATTTACTGAGGAGCTCTGAGGAGTCCATCTTGTGGAGATGATTCAG 480  
 QY 481 ATGCTGAGCGCATGAGGCTTACTTACTGCTGCTCAAGGAGCGTGATCATTCTTCTGT 540  
 Db 481 ATGCTGAGCGCATGAGGCTTACTTACTGCTGCTCAAGGAGCGTGATCATTCTTCTGT 540  
 QY 541 GAGGTACAGGCTTTATTTGAAGCTGAGGCTGTGAGACACTTCTTTTGAACCTCTCTC 600  
 Db 541 GAGGTACAGGCTTTATTTGAAGCTGAGGCTGTGAGACACTTCTTTTGAACCTCTCTC 600  
 QY 541 GAGGTACAGGCTTTATTTGAAGCTGAGGCTGTGAGACACTTCTTTTGAACCTCTCTC 600  
 Db 541 GAGGTACAGGCTTTATTTGAAGCTGAGGCTGTGAGACACTTCTTTTGAACCTCTCTC 600  
 QY 601 TTTGCTGCTGCTGCT 660  
 Db 601 TTTGCTGCTGCTGCT 660  
 QY 661 TGCATCTAGGCGCTGCT 720  
 Db 661 TGCATCTAGGCGCTGCT 720  
 QY 721 ACCTGCT 780  
 Db 721 ACCTGCT 780  
 QY 781 CTGAGGCGCTAGGCGCT 840  
 Db 781 CTGAGGCGCTAGGCGCT 840  
 QY 841 GTCCTTACT 900  
 Db 841 GTCCTTACT 900  
 QY 901 GCACTGAGGAGGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCG 948  
 Db 901 GCACTGAGGAGGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCG 948  
 RESULT 9  
 ABK37582  
 ID ABK37582 standard; cDNA; 948 BP.  
 XX  
 AC ABK37582;  
 XX  
 DT 08-MAY-2002. (first entry)  
 XX  
 DE DNA encoding G-coupled olfactory receptor #84.  
 XX  
 KW Human; olfactory G-coupled receptor; sensory perception of odourant;  
 odour composition; taste composition; gene; ss.  
 XX

OS Homo sapiens.  
 XX  
 PN WO200198526-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PE 22-JUN-2001; 2001WO-US20122.  
 XX  
 PR 22-JUN-2000; 2000US-213812P.  
 PR 13-MAR-2001; 2001US-0804291.  
 XX  
 PA (SENO-) SENOMYX INC.  
 XX  
 PI Zozulya S, Stryer L;  
 XX  
 DR WPI; 2002-083330/11.  
 DR P-PSDB; AA085223.  
 XX  
 PT Representing sensory perception of one or more odourants for the  
 PT identification and design of tastes and odours comprises providing a  
 PT representative group of n olfactory receptors -  
 XX  
 PS Example; Page 96; 182pp; English.  
 XX  
 CC The invention relates to a method of representing sensory perception of  
 CC one or more odourants. The method comprises: (a) providing a  
 CC representative class of n olfactory receptors or ligand binding domains  
 CC (LBD) of these receptors; (b) measuring values X1 to Xn representative  
 CC of at least one activity of one or more odourants selected from:  
 CC (i) binding one or more odourants to the LBD of at least one of the n  
 CC olfactory receptors; (ii) activating at least one of the n  
 CC olfactory receptors with the one or more odourants; and (iii) blocking at  
 CC least one of the n olfactory receptors with the one or more odourants;  
 CC and (c) generating a representation of sensory perception from the values  
 CC X1 to Xn. The representation of the sensory perception of odourants is  
 CC useful for the design and formulation of odour and taste compositions.  
 CC ABK37499-ABK37754 and ABK37918-ABK37921 represent human olfactory  
 CC G-coupled receptor coding sequences and related PCR primers of the  
 CC invention.  
 XX  
 SQ Sequence 948 BP; 176 A; 279 C; 228 G; 265 T; 0 other;  
 Query Match 85.5%; Score 810.4; DB 24; Length 948;  
 Best Local Similarity 90.9%; Pred. No. 2.3e-217;  
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
 QY 1 ATGGAAGATGGGTGAACAGTCTCTACACAGATGGCTTCTCTCTTGGGATCTTTCC 60  
 Db 1 ATGGAAGATGGGTGAACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
 QY 61 CACAGCGAGCTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 Db 61 CACAGTACTGCTGACCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 121 TGTGGGAATGTCCT 180  
 Db 121 TGTGGGAATGTCCT 180  
 QY 181 TACTTCTTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 Db 181 TACTTCTTCTCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 QY 241 AAGATGGCAGCACT 300  
 Db 241 AAGATGGCAGCACT 300  
 QY 301 CAAATTTGGCTTTTGTCT 360  
 Db 301 CAAATTTGGCTTTTGTCT 360  
 QY 361 TATGACCACTAGCTGAGCGCTTACCCACCTCATCTTACCTCATGAAATCAGAGG 420  
 Db 361 TATGACCGCTATGCTGAGCGCTTACCCACCTCATCTTACCTCATGAAATCAGAGG 420

```
QY 421 GTCGTCTCCAGATTACTGGAGCTCTGGCCCTTGGGATAATAGATGAGTGTATCAG 480
    |||||
DB 421 GTCGTCTCCAGATTACTGGAGCTCTGGCCCTTGGGATAATGATGGCTGTATCCAG 480
QY 481 ATGCTGGCAGCCATGGGCTTACTTACTGTGGCTCAAGGAGCGTGATCATTCTTCTGT 540
    |||||
DB 481 ATGCTGGTAGTATGTAATTTCCCTACTGTGGCTTGAGGAAGGTGAACCATTTCTCTGT 540
QY 541 GAGTACAGCTTATTTGAAGCTGGCTGTGCAGACACTTCCCTTTTGACACCCCTCCTC 600
    |||||
DB 541 GAGATGCTATCCTTGTGTAAGCTGGCTGTGTAACACATCCCTGTGAGAAGGTGATA 600
QY 601 TTTCCTTGTCTGTCTCATGCTTCTCCTTCTCCCTTCCATCATGAGCCCTCATGCT 660
    |||||
DB 601 TTTCCTTGTCTGTCTCATGCTTCTCCTTCTCCCTTCCATCATGAGCCCTCATGCT 660
QY 661 TGCATCCTAGGGGCTGTGCTCCGATATAGCTTGTCTCAGGCTGGAAGGAGCCCTGGCC 720
    |||||
DB 661 CACATTTCTAGGAGCTGTGCTGCAATGACACTGTGCTCAGGCTGGAAGGAGCCCTGGCC 720
QY 721 ACCGTCTCTCCACCTACAGCTGTACCCCTTCTATGAGGAGCCATGTTCATGTAC 780
    |||||
DB 721 ACCGTCTCTCCACCTACAGCTGTACCCCTTCTATGAGGAGCCATGTTCATGTAC 780
QY 781 CTGAGGCGCTAGGCGCTACCGGCGCTAGCCATGACAAGTGCGCTCTATCTTACACA 840
    |||||
DB 781 CTGAGGCGCTAGGCGCTACCGGCGCTAGCCATGACAAGTGCGCTCTATCTTACACA 840
QY 841 GTCCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGGG 900
    |||||
DB 841 GTCCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGGG 900
QY 901 GCACTGAGGAAGGCGCTGACCGCTGCAGATTGGCAGCCAGCACTGA 948
    |||||
DB 901 GCACTGAGGAAGGCGCTGACCGCTGCAGATTGGCAGCCAGCACTGA 948

RESULT 10
AAH31875
ID AAH31875 standard; DNA; 945 BP.
XX
AC AAH31875;
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 448.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA ) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
XX sensation for identifying olfactory agonists and antagonists -
```

```
PS Claim 8; Page 359-360; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 945 BP; 174 A; 279 C; 228 G; 264 T; 0 other;
Query Match 85.3%; Score 809; DB 22; Length 945;
Best Local Similarity 91.0%; Pred. No. 5.6e-217;
Matches 860; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 1 ATGGGAAGATGGGTGAACCAAGTCTACACAGATGGCTTCTCTTGGCATCTTTCC 60
    |||||
DB 1 ATGGAGACGTGGGTGAACCAAGTCTACACAGATGGCTTCTCTTGGCATCTTTCC 60
QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTGACATATAGTGTGCTTCACAGTGGCCCTC 120
    |||||
DB 61 CACAGTACTGTGACCTTGTCTCTCTCTCTGATGAGGCTTTCACAGTGGCCCTC 120
QY 121 TGTGGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
    |||||
DB 121 TGTGGGAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 TACTTCTCTCAGCCAGCTCTCCCTCATATGACCTCATATGTTGCTGTAACTGTGCCA 240
    |||||
DB 181 TACTTCTCTCAGCCAGCTCTCCCTCATATGACCTCATATGTTGCTGTAACTGTGCCA 240
QY 241 AAGATGGCAGCCACTTCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATA 300
    |||||
DB 241 AAGATGGCAGCCACTTCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATA 300
QY 301 CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTGCTGGACTATGGCT 360
    |||||
DB 301 CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTGCTGGACTATGGCT 360
QY 361 TATGACCACTAGCGGCGCTTAGCCACCCACTTCACTATCCCATCCCATGAATCAGAG 420
    |||||
DB 361 TATGACCGCTATGCGGCGCTTAGCCACCCACTTCACTATCCCATGAATCAGAG 420
QY 421 GTCGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATAATAGATGAGTGTATCAG 480
    |||||
DB 421 GTCGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATAATGATGGCTGTATCCAG 480
QY 481 ATGCTGGCAGCCATGGGCTTACTTACTGTGGCTCAAGGAGCGTGATCATTCTCTGT 540
    |||||
DB 481 ATGCTGGTAGTATGTAATTTCCCTACTGTGGCTTGAGGAAGGTGAACCATTTCTCTGT 540
QY 541 GAGTACAGCTTATTTGAAGCTGGCTGTGCAGACACTTCCCTTTTGACACCCCTCCTC 600
    |||||
DB 541 GAGATGCTATCCTTGTGTAAGCTGGCTGTGTAACACATCCCTGTTGAGAAGGTGATA 600
QY 601 TTTCCTTGTCTGTCTCATGCTTCTCCTTCTCCCTTCCATCATGAGCCCTCATGCT 660
    |||||
DB 601 TTTCCTTGTCTGTCTCATGCTTCTCCTTCTCCCTTCCATCATGAGCCCTCATGCT 660
QY 661 TGCATCCTAGGGGCTGTGCTCCGATATAGCTTGTCTCAGGCTGGAAGGAGCCCTGGCC 720
    |||||
DB 661 CGCATTTAGGAGCTGTGCTCAATGACACTTGTCTCAGGCTGGAAGGAGCCCTGGCC 720
QY 721 ACCGTCTCTCCACCTAACAGCTGTACCCCTTCTATGAGGAGCCATGTTCATGTAC 780
    |||||
```



Db 721 ACCTGCTCCACCTGACAGCTGTACACCCTCTCTATAGGGGACGCCATGTTCACTAC 780  
QY 781 CTGAGCCCTAGGCGCTACCGGCGCTAGCCATGACAGAGTGCGCTCTATCTTACACA 840  
Db 781 CTGAGCCCTAGGCGCTACCGGCGCGCCAGCCATGACAGAGTGCGCTCTATCTTACACG 840  
QY 841 GTCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGAAATGGAGTGATGGG 900  
Db 841 GTCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGAAATGGAGTGATGGG 900  
QY 901 GCACGTAGGAAGGGGCTGAGCGCTGACAGATTGGCAGCCAGCAC 945  
Db 901 GCACGTAGGAAGGGGCTGAGCGCTGACAGATTGGCAGCCAGCAC 945

## RESULT 11

AAH31968  
ID AAH31968 standard; DNA; 945 BP.  
XX  
AC AAH31968;  
XX  
DT 30-JUL-2001 (first entry)  
XX  
DE Human olfactory receptor polynucleotide, SEQ ID NO: 541.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200127158-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 06-OCT-2000; 2000WO-US27582.  
XX  
PR 08-OCT-1999; 99US-0158615.  
PR 24-FEB-2000; 2000US-0184809.  
XX  
PA (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX  
DR WPI; 2001-290713/30.  
XX  
PT New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX  
PS Claim 8; Page 396-397; 1857pp; English.  
XX  
CC The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.  
XX  
SQ Sequence 945 BP; 175 A; 279 C; 227 G; 264 T; 0 other;

Query Match 85.2%; Score 807.4; DB 22; Length 945;  
Best Local Similarity 90.9%; Pred. No. 1.6e-216;  
Matches 859; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 ATGGAAGATGGGTGAACACAGTCCCTACACAGATGGCTTCTTCCCTTGGCATCTTTCC 60  
Db 1 ATGGAAGATGGGTGAACACAGTCCCTACACAGATGGCTTCTTCCCTTAGCATCTTCTCC 60  
QY 61 CACAGCCAGACATGACCTTGTCTCTCTCTGACGTATATGTTGTTCTTACAGTGGCCCTC 120  
Db 61 CACAGTACTGCTGACCTTGTCTCTCTCTCTCCGTGTTATGGCGTTCTTACAGTGGCCCTC 120  
QY 121 TGTGGGAATGTCTCTCTCATATCTTCTCATCTTACCTGAGCGCTGAGACTTCACACCCCATG 180  
Db 121 TGTGGGAATGTCTCTCTCATATCTTCTCATCTTACATGAGACCTTCACACCCCATG 180  
QY 181 TACTTCTTCCATGAGCCAGCTCTCCCTCATGAGCTCATGTGTGTTCTGTACCAATGTGCCA 240  
Db 181 TACTTCTTCCATGAGCCAGCTCTCCCTCATGAGCTCATGTGTGTTCTGTACCAATGTGCCA 240  
QY 241 AAGATGGCAGCCCACTTCTCTGTGGCAGAGAGTCCATCTCTTGTGGGCTGTGGCATA 300  
Db 241 AAGATGGCAGCCCACTTCTCTGTGGCAGAGAGTCCATCTCTTGTGGGCTGTGGCATA 300  
QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTTGCTGGGACTCATGGCT 360  
Db 301 CAAATTTGGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTTGCTGGGACTCATGGCT 360  
QY 361 TATGACCACTAGCTGGCGCTTACCTTACCTTACTGTGGCTCAAGAGCGGTGATCATTCTTCTGT 420  
Db 361 TATGACCGCTATGTGGCCATTTAGCCACCCACTTCACTATCCCATCTCATGATGATCAGAG 420  
QY 421 GTCTGTCTCCAGATTACTGGGAGCTCTTGGGCTTTGGGATATAGATGGAGTGAATTCA 480  
Db 421 GTCTGTCTCCAGATTACTGGGAGCTCTTGGGCTTTGGGATATAGATGGGCTTGAATCCAG 480  
QY 481 ATGTTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCATTCTTCTGT 540  
Db 481 ATGTTGGTAGTATGAATTTCCCTTACTGTGGCTTGAAGGATGAACCATTTCTTCTGT 540  
QY 541 GAGCTACAAGCTTTATGAGCTGGGCTGTGACAGACACTTCCCTTTTGGACACCCCTCCTC 600  
Db 541 GAGATGCTATCTTGTGTAAGCTGGGCTGTGTAAGACACATCCCTGTTTGAAGAGTGTATA 600  
QY 601 TTGCTGTCTGTCTTCTCATGCTTCTCTTCCCTTCTCCATCATCATGCTGCTTCTATGCT 660  
Db 601 TTGCTGTCTGTCTTCTCATGCTTCTCTTCCATCTTCCATCATCATGCTGCTTCTATGCT 660  
QY 661 TGCATCTAGGGGCTGTGCTCCGAATAGCGCTGTGCTCAGGCGCTGGAAGGCGCTGGCC 720  
Db 661 CACATTTAGGAGCTGTGCTCGAATGACACTGTGCTCAGGCGCTGGAAGGCGCTGGCC 720  
QY 721 ACCTGCTCTCCACCTAATACAGCTGTCAACCTCTTCTATGAGGCGAGCCATGTTCACTGAC 780  
Db 721 ACCTGCTCTCCACCTGACAGCTGTCAACCTCTTCTATGAGGCGAGCCATGTTCACTGAC 780  
QY 781 CTGAGGCTTAGGCGCTACCGGCGCTAGCCATGACAGAGTGCGCTCTATCTTACACA 840  
Db 781 CTGAGGCTTAGGCGCTACCGGCGCGCCAGCCATGACAGAGTGCGCTCTATCTTACACG 840  
QY 841 GTCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGAAATGGAGTGATGGG 900  
Db 841 GTCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGAAATGGAGTGATGGG 900  
QY 901 GCACGTAGGAAGGGGCTGAGCGCTGACAGATTGGCAGCCAGCAC 945  
Db 901 GCACGTAGGAAGGGGCTGAGCGCTGACAGATTGGCAGCCAGCAC 945

## RESULT 12

AAH31968  
ID AAD26374 standard; cDNA; 965 BP.  
XX  
AC AAD26374;  
XX  
DT 26-MAR-2002 (first entry)  
XX

DE	Human G-protein coupled receptor 6 (GCRC-6) cDNA.	
XX	Human; G-protein coupled receptor 6; cell proliferative disorder;	
KW	arteriosclerosis; hepatitis; cancer; neurological disorder; epilepsy;	
KW	Alzheimer's disease; Parkinson's disease; cardiovascular disorder;	
KW	atherosclerosis; hypertension; myocardial infarction; peptic ulcer;	
KW	gastrointestinal disorder; dysphagia; anorexia; autoimmune disorder;	
KW	acquired immune deficiency syndrome; inflammatory disorder; infection;	
KW	Addison's disease; allergy; Grave's disease; metabolic disorder; AIDS;	
KW	diabetes; obesity; osteoporosis; gene therapy; GCRC-6; ss.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..816
FT		/*tag= a
FT		/product= "Human GCRC-6 protein"
XX		
PN	MO200187937-A2.	
XX		
PD	22-NOV-2001.	
XX		
PF	17-MAY-2001; 2001WO-0516285.	
XX		
PR	18-MAY-2000; 2000US-205628P.	
PR	22-MAY-2000; 2000US-206222P.	
PR	25-MAY-2000; 2000US-207566P.	
PR	02-JUN-2000; 2000US-208834P.	
PR	02-JUN-2000; 2000US-208861P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Patterson C, Lu DAM, Thornton M, Lu Y, Tribouley CM, Graul R;	
PI	Khan FA, Gandhi AR, Walla NK, Nguyen DB, Yue H, Hafalia A;	
PI	Elliott VS, Lal P, Reddy R, Kallick DA, Tang TY, Au-Young J;	
XX		
DR	MP1: 2002-089844/12.	
XX	P-PSDB; AAE16175.	
PT	Novel G-protein coupled receptors and polynucleotides useful for	
PT	diagnosis, treatment and prevention of disorders of cell proliferation,	
PT	neurological, cardiovascular, metabolic disorders and viral infections	
PT	-	
PS	Claim 5; Page 114; 115pp; English.	
XX		
CC	The invention relates to human G-protein coupled receptor (GCRC)	
CC	polypeptides and polynucleotides. GCRC polypeptides are useful for	
CC	screening compounds that modulate their activity. They are useful in	
CC	the diagnosis, prevention and treatment of disorders which include	
CC	cell proliferative disorders such as arteriosclerosis, hepatitis,	
CC	myelofibrosis, psoriasis and cancer including adenocarcinoma, leukaemia	
CC	lymphoma; neurological disorders such as epilepsy, ischaemic	
CC	cerebrovascular disease, Alzheimer's disease, Pick's disease, dementia,	
CC	Parkinson's disease, ataxias, multiple sclerosis, bacterial and viral	
CC	meningitis, Creutzfeldt-Jakob disease, schizophrenic disorders, amnesia	
CC	cardiovascular disorders such as arteriovenous fistula, atherosclerosis	
CC	hypertension, vascular tumours, myocardial infarction, hypertensive	
CC	heart disease, infective endocarditis, cardiomyopathy, myocarditis;	
CC	gastrointestinal disorders such as dysphagia, peptic oesophagitis,	
CC	emesis, anorexia, nausea, peptic ulcer, cholelithiasis, diarrhoea,	
CC	constipation, acquired immune deficiency syndrome (AIDS), hepatic	
CC	encephalopathy; autoimmune/inflammatory disorders such as Addison's	
CC	disease, allergies, spondylitis, amyloidosis, anaemia, asthma, contact	
CC	dermatitis, Crohn's disease, diabetes mellitus, Goodpasture's syndrome,	
CC	emphysema, Grave's disease, gout, multiple sclerosis, rheumatoid	
CC	arthritis, systemic lupus erythematosus, uveitis, viral, bacterial,	
CC	fungal, parasitic, protozoal and helminthic infections and trauma;	
CC	metabolic disorders such as diabetes, obesity and osteoporosis; and	
CC	viral infections such as infection caused by viral agent classified as	
CC	adenovirus, arenavirus, bunyavirus. Polynucleotides of the invention	
CC	are useful as probes for assessing toxicity of test compounds. They are	
CC	also used in gene therapy. The present sequence is human G-protein	

ID	Accession	Standard	CDNA	BP
AC	AAA51179	standard	CDNA	669 BP
DT	AAA51179			
DE	26-SEP-2000	(first entry)		
KW	Murine olfactory receptor ligand-binding region	cdna 5.		
KW	Degenerate primer; ligand-binding region; olfactory receptor;			
KW	Transmembrane domain; N-terminal; plasma membrane; translocation domain;			
KW	human rhodopsin receptor; odorant; toxicity; olfactory response; ss.			
OS	Mus musculus.			
FT	Key	Location/Qualifiers		
FT	CDS	1..669		
FT		/*tag= a		
FT		/product= Olfactory_receptor_ligand_binding_protein		
FT		/partial		
PN	WO200035274-A1.			
XX				
PD	22-JUN-2000.			
XX				
PF	17-DEC-1999;	99WO-US30221.		
XX				
PR	17-DEC-1998;	98US-0112605.		
XX				
PA	(UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.			
XX				
PI	Reed RR, Krautwurst D, Yau KW;			
XX				
DR	WPI; 2000-431471/37.			
DR	P-PSDB; AAY96666.			
PT	Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.			
PT	for producing receptor libraries used in e.g. screening odorants for			
PT	toxicity			
XX				
PS	Claim 18; Page 31-32; 61pp; English.			
XX				
CC	AAA51175-93 encode murine odorant/ligand binding regions of olfactory			
CC	receptors. They were amplified using degenerate primers shown in			
CC	AAA51163-64.			
CC	The primers can be used to generate a library of OR-LBR comprising the			
CC	transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric			
CC	nucleic acids encoding a 7-TMD protein comprise a sequence encoding an			
CC	N-terminal plasma membrane translocation domain (especially the sequence			
CC	shown in AAY96658), a first TMD and the OR-LBR. The translocation domain			
CC	was initially derived from the N-terminus of the human rhodopsin			
CC	receptor. LBR produced by amplification with the primers are useful for			
CC	generating new odorants, to screen for toxicity or therapeutic activity			
CC	in odorants, and altering an animal's olfactory response (claimed).			
XX				
SQ	Sequence 669-BP; 124 A; 190 C; 148 G; 207 T; 0 other;			
QY	Query Match	55.5%;	Score 526.6;	DB 21; Length 669;
QY	Best Local Similarity	86.7%;	Pred. No. 9.6e-138;	
QY	Matches 580; Conservative	0;	Mismatches 89;	Indels 0; Gaps 0;
DB	1	AGTCAGCTCTCCCTCATGAGACCTCATGCTGTGTAACATTGTGCCAAGATGCGAGTC	60	
QY	193	AGCCAGCTCTCCCTCATGAGACCTCATGCTGTGTAACATTGTGCCAAGATGCGAGTC	252	
DB	1	AGTCAGCTCTCCCTCATGAGACCTCATGCTGTGTAACATTGTGCCAAGATGCGAGTC	60	
QY	253	AACCTCTCTGCTGGCAGGAAGTCCATCTCTTTGGGGCTGTGGCATACAAATTGGCTTT	312	
DB	61	AACCTCTCTGCTGGCAGGAAGTCCATCTCTTTGGGGCTGTGGCATACAAATTGGCTTT	120	
QY	313	TTTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGACCACTAC	372	
DB	121	TTTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGACCACTAT	180	

Accession	Gene	Species	Length (bp)	Location	Qualifiers
QY 373	GTGGCCGTTAGCCACCCACTTCACTATATCCCATCTCATGATCAGAGGGTCTGTCACG	Human	432		
Db 181	GTGGCCATTAGCCACCCACTTCACTATATCCCATCTCATGATCAGAGGGTCTGTCACG	Human	240		
QY 433	ATTACTGGAGCTCCTGGGCTTTGGGATTAATAGATGAGTGTTCAGATGTGGCAGCC	Human	492		
Db 241	ATTGCTGGAAGTTCCTGGGCTTTGGGATTCCTTGATGGAATAATTCAGATGGTGCCAGCC	Human	300		
QY 493	ATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATATCTTTTCTGTGAGGTACAGCT	Human	552		
Db 301	ATGAGCCCTGCTTACTGTGGCTCAAGCGTATATAGATCACTTCTTCTGTGAAGTGCCGCT	Human	360		
QY 553	TTATGAGCTGGGCTGTGACAGACACTTCCCTTTTGAACACCCCTCTCTTGTGCTGT	Human	612		
Db 361	TTACTGAAGCTGGGCTGTGACAGACACTTCCCTTTTGAACACCCCTCTCTTGTGCTGT	Human	420		
QY 613	GCTTTCATGCTTCT	Human	672		
Db 421	GCTTTCATGCT	Human	480		
QY 673	GCTGTGCTCCGATACGCTCTGCTCAAGCCCTGGAAGGAGCCCTGGCCACTGCTCTCC	Human	732		
Db 481	GCTGTGCTCCGATACGCTCTGCTCAAGCCCTGGAAGGAGCCCTGGCCACTGCTCTCC	Human	540		
QY 733	CACCTAAGCTGTACACCT	Human	792		
Db 541	CACCTAAGCTGTACACCT	Human	600		
QY 793	CGCTACGCGGCGGCTAGCCATGACAGGTGGCTCTATCTTCTACAGTCTTACTCTCC	Human	852		
Db 601	CGATATGCGGCT	Human	660		
QY 853	ATGCTGAC 861	Human			
Db 661	ATGCTGAC 669	Human			
RESULT 14					
ID	AAA51183	standard; cdna; 669 BP.			
XX AC	AAA51183;				
XX DT	26-SEP-2000 (first entry)				
XX DE	Murine olfactory receptor ligand-binding region cdna 9.				
XX KM	Degenerate primer; ligand-binding region; olfactory receptor;				
KW	Transmembrane domain; N-terminal; plasma membrane; translocation domain;				
KW	human rhodopsin receptor; odorant; toxicity; olfactory response; ss.				
XX OS	Mus musculus.				
XX FH	Key	Location/Qualifiers			
FT CDS	1..669				
FT FT	/*tag= a				
FT FT	/product= Olfactory_receptor_ligand_binding_protein				
FT FT	/partial				
XX PN	W0200035274-A1.				
XX PD	22-JUN-2000.				
XX PF	17-DEC-1999; 99WO-US30221.				
XX PR	17-DEC-1998; 98US-0112605.				
XX PA	(UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.				
XX PI	Reed RR, Krautwurst D, Yau KW;				
XX DR	WPI: 2000-431471/37.				

DR P-PSDB; AAY96670.  
 XX  
 PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.  
 PT for producing receptor libraries used in e.g. screening odorants for  
 PT toxicity  
 XX  
 PS Claim 18; Page 33; 61pp; English.  
 CC  
 CC AAA51175-93 encode murine odorant/ligand binding regions of olfactory  
 CC receptors. They were amplified using degenerate primers shown in  
 CC AAA51163-64.  
 CC The primers can be used to generate a library of OR-LBR comprising the  
 CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric  
 CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an  
 CC N-terminal plasma membrane translocation domain (especially the sequence  
 CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain  
 CC was initially derived from the N-terminus of the human rhodopsin  
 CC receptor. LBR produced by amplification with the primers are useful for  
 CC generating new odorants, to screen for toxicity or therapeutic activity  
 CC in odorants, and altering an animal's olfactory response (claimed).  
 XX  
 SQ Sequence 669 BP; 124 A; 190 C; 148 G; 207 T; 0 other:  
 Query Match 55.5%; Score 526.6; DB 21; Length 669;  
 Best Local Similarity 86.7%; Pred. No. 9.6e-138;  
 Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
 QY 193 AGCCAGCTCTCCCTCAGGACCTCAGTGTGGTCTGTACATTTGCCAAGATGGCAGCC 252  
 DB 1 AGTCAGCTCTCCCTCAGGACCTCAGTGTGGTCTGTACATTTGCCAAGATGGCAGTC 60  
 QY 253 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATTGGCTTT 312  
 DB 61 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATCGGATTT 120  
 QY 313 TTGTCTCTCTGTGGGATCTGAGGGGCTCTGTGGGACTCATGGCTTATGACCACTAC 372  
 DB 121 TTGTCTCTCTGTGGGATCTGAGGGGCTCTGTGGGACTCATGGCTTATGATCGCTAT 180  
 QY 373 GTGGCCGTTAGCCACCACCTTCACATTCCTCATGATGAGAGGCTGTCTCAG 432  
 DB 181 GTGGCCATTTAGCCACCACCTTCACATTCCTCATGAGAGGCTGTCTCAG 240  
 QY 433 ATTACTGGGAGCTCCTGGGCTTTGGGATATAGATGAGTGTGATGATGGTGGCAGCC 492  
 DB 241 ATTGCTGGAAGTCTGGGCTTTGGGATCTTGAATGATGATGATGGTGGCAGCC 300  
 QY 493 ATGGGCTTACCTTACTGTGGCTCAAGAGGCTGTGATCACTTTTCTGTGAGGTACAGCT 552  
 DB 301 ATGAGCTGCCCCCTACTGTGGCTCAGCGTATATAGATCACTTCTGTGAAGTGGCGGCT 360  
 QY 553 TTATTTGAGGCTGGGCTGTGAGACACTTCCCTTTTGAACACCCTCTCTTGTGCTGTGT 612  
 DB 361 TTACTGAGGCTGGGCTGTGAGACACTTCCCTTTTGAACACCCTCTCTTGTGCTGTGT 420  
 QY 613 GTCTTATGCTTCTCTCTCTCTCTCTCATCATGAGGCTCTCATGCTTGCATCTCAGGG 672  
 DB 421 GTCTTATGCTTCTCTCTCTCTCTCTCATCATGAGGCTCTCATGCTTGCATCTCAGGG 480  
 QY 673 GCTGTGCTCGAATAGCTGTGCTCAGAGGCTGGAAGAAAGCCCTGGCCACCTGCTCTCC 732  
 DB 481 GCTGTGCTCGAATAGCTGTGCTCAGAGGCTGGAAGAAAGCCCTGGCCACCTGCTCTCC 540  
 QY 733 CACCTTACAGCTGTACACCTCTTCTATGAGGGGAGCCAGTTCATGCTAGGAGGCTTACG 792  
 DB 541 CACCTTACAGCTGTCTCTCTCTCTCTCTACAGGGGAGCAATGTTCATCTACCTGAGGCAAG 600  
 QY 793 CGCTACGGGGCCCTAGCCATGACAAAGTGGCTCTATCTTCTACACAGTCTTACTTCC 852  
 DB 601 CGATATGCGGCTCTAGCCATGACAAAGTGTCTCAATCTTCTACACAGTCTTACTTCC 660  
 QY 853 ATGCTGAAC 861  
 11111 111

DB 661 ATGCTCAAC 669  
 RESULT 15  
 ID AAA51181  
 ID AAA51181 standard; cDNA; 669 BP.  
 AC AAA51181;  
 DT 26-SEP-2000 (first entry)  
 DE Murine olfactory receptor ligand-binding region cDNA 7.  
 KW Degenerate primer; ligand-binding region; olfactory receptor;  
 KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;  
 KW human rhodopsin receptor; odorant; toxicity; olfactory response; ss.  
 OS Mus musculus.  
 FH key Location/Qualifiers  
 FT CDS . 1..669  
 FT /\*tag= a  
 FT /product= Olfactory\_receptor\_ligand\_binding\_protein  
 FT /partial  
 PN WO200035274-A1.  
 PD 22-JUN-2000.  
 PF 17-DEC-1999; 99WO-US30221.  
 PR 17-DEC-1998; 98US-0112605.  
 PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 PI Reed RR, Krautwurst D, Yau KW;  
 XX WPI; 2000-431471/37.  
 DR P-PSDB; AAY96668.  
 PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.  
 PT for producing receptor libraries used in e.g. screening odorants for  
 PT toxicity  
 XX  
 PS Claim 18; Page 32; 61pp; English.  
 CC  
 CC AAA51175-93 encode murine odorant/ligand binding regions of olfactory  
 CC receptors. They were amplified using degenerate primers shown in  
 CC AAA51163-64.  
 CC The primers can be used to generate a library of OR-LBR comprising the  
 CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric  
 CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an  
 CC N-terminal plasma membrane translocation domain (especially the sequence  
 CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain  
 CC was initially derived from the N-terminus of the human rhodopsin  
 CC receptor. LBR produced by amplification with the primers are useful for  
 CC generating new odorants, to screen for toxicity or therapeutic activity  
 CC in odorants, and altering an animal's olfactory response (claimed).  
 XX  
 SQ Sequence 669 BP; 125 A; 190 C; 147 G; 207 T; 0 other:  
 Query Match 55.4%; Score 525; DB 21; Length 669;  
 Best Local Similarity 86.5%; Pred. No. 2.7e-137;  
 Matches 579; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
 QY 193 AGCCAGCTCTCCCTCAGGACCTCAGTGTGGTCTGTACATTTGCCAAGATGGCAGCC 252  
 DB 1 AGTCAGCTCTCCCTCAGGACCTCAGTGTGGTCTGTACATTTGCCAAGATGGCAGTC 60  
 QY 253 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATTGGCTTT 312  
 DB 61 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATCGGATTT 120

QY 313 TTGTCTCTCTTTGGGATCTGAGGGCTCTTGCTGGGACTCATGGCTTATGACCACTAC 372  
 |||||  
 Db 121 TTGTCTCTCTTTGGGATCTGAGGGCTCTTGCTGGGACTCATGGCTTATGACCACTAT 180  
 QY 373 GTGGCCGTAGCCACCACCTTCACCTATCCCATCCATGAATCAGAGGGTCTGTCCAG 432  
 |||||  
 Db 181 GTGGCCATTAGCCACCACCTTCACCTATCCCATCTCATGAGCCAAAGGTCTGTCCAG 240  
 QY 433 ATTACTGGAGCTCTGGGCTTTGGGATAATAGATGAGATTCAGATGTTGGCAGCC 492  
 |||||  
 Db 241 ATTGCTGAAGTTCCTGGGCTTTGGGATCCCTGATGAATAATTCAGATGTTGGCAGCC 300  
 QY 493 ATGGGCTTACCTTACTGTGGCTCAAGAGCGTGAATCACCCTTTCTGTGAGGTACAAGCT 552  
 |||||  
 Db 301 ATGAGCCTGCCCCACTGTGGCTCAGCGTATATAGATCACCCTTCTGTGAGTGGCGCT 360  
 QY 553 TTATTGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGTGT 612  
 |||||  
 Db 361 TTACTGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGTGT 420  
 QY 613 GTCTTCATGCTTCTCCTTCCCTTCTCCATCATATGGCCCTCCATGTGCTGATGCTAGGG 672  
 |||||  
 Db 421 GTCTTATGCTGCTCTCTCTTCTCGATCATGTGACTTCCATGCTCGCATCTTGGGG 480  
 QY 673 GCTGTGCTCCGAAATACGCTCTGTCTCAGGCCCTGGAAAAAGCCCTGGCCACCTGCTCCTCC 732  
 |||||  
 Db 481 ACTGTGCTCCGTAATGCACTCTGCCCCAGTCCCGAAAAAGGCCCTGGCCACTTGTTCCTCC 540  
 QY 733 CACCTAACAGCTGTCAACCTCTCTATGAGGGCAGCCATGTTCATGTACCTGAGGCCCTAGG 792  
 |||||  
 Db 541 CACCTGACAGCTGTCTCTCTCTACGGGGCAGCAATGTTCATCTACCTGAGGCCAAGG 600  
 QY 793 CGCTACCGGGGCCCTAGCCATGACAAGGTGGCCCTATCTTACACAGTCCCTTACTCCC 852  
 |||||  
 Db 601 CGATATCGCGGCTCCTAGCCATGACAAGTGTCTCAATCTTCTACACAGTCTTACTCCT 660  
 QY 853 ATGCTGAAC 861  
 |||||  
 Db 661 ATGCTCAAC 669

Search completed: May 9, 2003, 03:37:26  
 Job time : 272 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2003, 02:56:25 ; Search time 1522 Seconds  
(without alignments)  
10087.596 Million cell updates/sec

Title: US-09-634-109D-1  
Perfect score: 948  
Sequence: 1 atgggaagatgggtgaacca.....ggattggcagccagcactga 948

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_trod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	380.6	40.1	398	17	AQ153969	AQ153969 HS_2250_A
2	327.8	34.6	664	10	BB538338	BB538338 BB538338
3	297	31.3	3063	11	AK016560	AK016560 Mus muscu
4	295.4	31.2	1394	11	AK017036	AK017036 Mus muscu
5	290	30.6	2021	11	BC016940	BC016940 Homo sapi
6	249.4	26.3	632	17	AZ765752	AZ765752 IM0562E19

c 7	221.4	23.4	650	17	AG067024	AG067024 Pan trogl
c 8	205	21.6	592	17	AQ055849	AQ055849 HS_2082_B
c 9	200.6	21.2	448	17	AQ425578	AQ425578 CITBI-El
10	198	20.9	915	14	BQ888002	BQ888002 AGENCOURT
11	195	20.6	518	13	BI912805	BI912805 603176272
12	187.4	19.8	1046	14	BQ892087	BQ892087 AGENCOURT
13	184.6	19.5	618	17	AZ868627	AZ868627 2M0180J22
14	183.2	19.3	796	12	BG197640	BG197640 RST17016
c 15	183	19.3	481	17	AQ970130	AQ970130 RPCI-23-3
c 16	181.8	19.2	698	17	BH069789	BH069789 RPCI-24-3
17	180.6	19.1	955	9	AL539245	AL539245 AL539245
18	178	18.8	642	17	AZ771031	AZ771031 1M0572L19
19	177.2	18.7	663	17	AG090554	AG090554 Pan trogl
20	176.8	18.6	788	17	BH111304	BH111304 RPCI-24-3
c 21	176.4	18.6	523	9	AI604386	AI604386 vv74C09.x
22	176	18.6	641	17	AZ787852	AZ787852 2M0034M04
23	172	18.1	853	17	BH331857	BH331857 CH230-46N
24	170.8	18.0	646	17	AZ638594	AZ638594 1M0498M06
25	170.6	18.0	680	17	AZ235331	AZ235331 RPCI-23-6
26	169.2	17.8	580	17	AZ418115	AZ418115 1M0194N03
c 27	169.2	17.8	692	17	AZ090606	AZ090606 RPCI-23-2
28	168.2	17.7	710	17	AG031161	AG031161 Pan trogl
c 29	167.4	17.7	628	17	AZ086388	AZ086388 RPCI-23-2
c 30	167	17.6	797	17	AZ909618	AZ909618 RPCI-24-2
31	166.4	17.6	401	17	AQ201950	AQ201950 RPCI11-46
c 32	165.8	17.5	580	10	BE543056	BE543056 601068985
c 33	165.6	17.5	629	17	AZ019257	AZ019257 RPCI-23-2
c 34	165	17.4	740	17	AZ607393	AZ607393 1M0429N15
c 35	164	17.3	638	9	AA681865	AA681865 vr44a12.S
c 36	162.8	17.2	656	17	BH078380	BH078380 RPCI-24-3
37	162.8	17.2	705	11	AF327904	AF327904 Homo sapi
38	162.8	17.2	751	12	BE856570	BE856570 7f64e08.x
39	162.4	17.1	1501	11	AK016338	AK016338 Mus muscu
40	162	17.1	619	9	AI148854	AI148854 qc70a02.x
41	160	16.9	678	17	BH067899	BH067899 RPCI-24-3
c 42	159.6	16.8	628	17	AZ977433	AZ977433 2M0253P17
c 43	159.4	16.8	507	17	AZ560017	AZ560017 RPCI-23-2
44	159.2	16.8	366	17	AZ896126	AZ896126 RPCI-24-1
c 45	158.2	16.7	522	17	AZ720443	AZ720443 RPCI-24-1

ALIGNMENTS

RESULT 1  
LOCUS AQ153969/c 398 bp DNA linear GSS 16-OCT-1998  
DEFINITION HS\_2250\_A1\_H04\_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2250 Col=7 Row=O, DNA sequence.

ACCESSION AQ153969  
VERSION AQ153969.1 GI:3546639  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
1 (bases 1 to 398)  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 2250 row: O column: 7

```
Class: BAC ends
High quality sequence stop: 398.
Location/Qualifiers
1. 398
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=2250 Col=7 Row=0"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      121 a      103 c      98 g      75 t      1 others
ORIGIN

Query Match      40.1%; Score 380.6; DB 17; Length 398;
Best Local Similarity 97.5%; Pred. No. 1.5e-87;
Matches 386; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 151 TACCTGACGCTGAGCTTCACACCCCATGTACTTCTCCAGCCAGCTCCCTCATG 210
      ||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 396 TACGTGAGCGTTGGACTTCACACCCCATGTAATTCTTCNTCAGCCAGCTCCCATCATG 337

QY 211 GACCTCATGTGGTCTGTAACTTGTGCCAAGATGGCAGCCAACTTCCTGTCTGGCAGG 270
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 336 GACATCATGTGGTCTGTAACTTGTGCCAAGATGGCAGCCAACTTCCTGTCTGGCAGG 277

QY 271 AAGTCCATCTCCTTTGTGGCGTGTGGCATACAATTGGCTTTTGTCTCTCTTGTGGGA 330
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 276 AAGTCCATCTCCTTTGTGGCGTGTGGCATACAATTGGCTTTTGTCTCTCTTGTGGGA 217

QY 331 TCTGAGGGGCTTGTCTGGGACTCATGGCTTATGACCACTAGCTGGCCGTTAGCCACCA 390
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 216 TCTGAGGGGCTTGTCTGGGACTCATGGCTTATGACCGCTAGCTGGCCGTTAGCCACCA 157

QY 391 CTTCACTATCCCATCTCTCATGATCAGAGGGTCTGTCTCCAGATTACTGGGAGCTCCTGG 450
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 156 CTTCACTATCCCATCTCTCATGATCAGAGGGTCTGTCTCCAGATTACTGGGAGCTCCTGG 97

QY 451 GCCTTTGGATAATAGATGAGATTCAGATGGTGGCAGCCATGGGCTTACTTACTGT 510
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 96 GCCTTTGGATAATAGATGAGATTCAGATGGTGGCAGCCATGGGCTTACTTACTGT 37

QY 511 GGCTCAAGAGCGGTGATCACTTTTCTGTGAGGTA 546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 36 GGCTCAAGAGCGGTGATCACTTTTCTGTGAGGTA 1

RESULT 2
BB538338 664 bp mRNA linear EST 26-OCT-2001
LOCUS BB538338 RIKEN full-length enriched, 0 day neonate eyeball Mus
DEFINITION musculus cDNA clone E130011E23 3', mRNA sequence.
ACCESSION BB538338
VERSION BB538338.2 GI:16446878
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 664)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT On Jul 31, 2000 this sequence version replaced gi:9593838.
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
```

```
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Location/Qualifiers
1. 664
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E130011E23"
/clone_lib="RIKEN full-length enriched, 0 day neonate
eyeball"
/tissue_type="eyeball"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGACGGCGCGCACTCGAGTTTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTTCTGAGTTAAATTAAATTAATCCCGCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLX I."
BASE COUNT      141 a      176 c      142 g      204 t      1 others
ORIGIN

Query Match      34.6%; Score 327.8; DB 10; Length 664;
Best Local Similarity 86.2%; Pred. No. 8.6e-74;
Matches 362; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 529 CACTTTTCTGTGAGGTACAGCTTTATGAAGCTGGCGCTGGCAGACACTCCCTTTT 588
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 1 CACTTCTTCTGTGAGGTGCGCGCTTTACGAAGCTGGCGCTGGCAGACACTCCCTTTTC 60

QY 589 GACACCGCTCTCTTGTGCTGTCTTCATGCTTCCTTCCCTTCATCATCATG 648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
```



```

/codon_start=1
/protein_id="BAB30304.1"
/db_xref="GI:1285358"
/db_xref="MGI:1921647"
/translation="MEPWNSTLESGFILVGLDGSPELLCATVTTLYMLALISNGL
LLLVITDARLHVPMYLLRLSLIDLFTSVTPNTVVDPLLRDNTISFEGALQLF
SAMTIGAEELLAFMAYDRVVAICHPLNMVIMSPKACRLMVAISWLASLSALGHT
VYTMHFPFCMSQEIIRHLICEVPLKLKACADTSQYELMAYVTGVIFLLPLSAIITSY
SLIFTVLHMPSENEGKRALVTCSSHLTVGMFYGGATFMYVLPSSEHSPKODNIISV
FYTIVTPALNPLIYSLRNEKIVIGAVRRVLRHILPAHAY"
3038..3043

```

```

polyA_signal
/note="putative"
3063
polyA_site
/note="putative"

```

```

BASE COUNT      879 a      612 c      696 g      876 t
ORIGIN

```

```

Query Match      31.3%; Score 297; DB 11; Length 3063;
Best Local Similarity 58.2%; Pred. No. 1.7e-65;
Matches 522; Conservative 0; Mismatches 375; Indels 0; Gaps 0;

```

```

QY 23 CCTACACAGATGGCTTCTTCTCTGGGCACTTTTCCACAGCCAGACTGACCTGTGC 82
    |||  |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 457 CCTGGAAGTGATTCATCTTGGTGGGATTCGTGATGGCAGTGGCTCTCTGAACTGC 516

QY 83 TCTTCTGCACTTATGGTGTCTTACAGATGGCCCTCTGTGGGAATGTCCTCATCT 142
    |||  |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 517 TCTGTGCCACAGTTACAACCCGTGTACATGTGGCACTGATCAGCAATGACTTCTACTCC 576

QY 143 TCCATCTACCTGGAGCGTGGACTTCACACCCCATGTACTTCTCTCAGCCAGCTCT 202
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 577 TGGTCATCAGATGGATGGCCGGCTTACAGTACCATGTACTCTCCTAGGACAGCTGT 636

QY 203 CCTCATGAGACTCATGTGGTCTGTACATGTGCCAAGATGGACCCACTTCTCTGT 262
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 637 CTCTCATGTACCTCTCTTACATCAGTTGTAACTCCCAACACTGTGTGATTTTCTGC 696

QY 263 CTGGCAGGAATCCATCTCCTTTGTGGGCTGTGGCATACAATTTGGCTTTTGTCTCTC 322
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 697 TCAGAGACAACCATATCTCTTGAAGGATGTGCCCTTCAATTTGTCTCAGCAATGACAT 756

QY 323 TTGTGGATCTGAGGGGCTCTTGTCTGGGACTATGGCTTATGACCACTACGTCGCCGTTA 382
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 757 TGGGTGTGACAGAGAGGCTCTTCTGTGGCTTATGAGTATGATGTGCCATTTT 816

QY 383 GCCACCCACTTCACTATCCCATCTCATGATCAGAGAGGCTCTCTCCAGATTAAGTGGGA 442
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 817 GTCATCTCTTAACATCATGATCTCATGATCCAAAGCCCTGACAGGCTCATGTGGCCA 876

QY 443 GCTCTGGGCTTTGGGATAATAGATGAGATTCAGATGCTGCGACAGCCATGGGCTTAC 502
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 877 TATCATGGATCTTACATCTCTTAGTGCCCTAGTTCACACAGTGTACACAATGACATTC 936

QY 503 CTTACTGTGGCTCAAGAGCGTGGATCATCTTTTCTGTGAGGTACAGCTTTTATGAAGC 562
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 937 CTTTGTGATGTCGCCAGGAATGACACACCTGCTCTGTGAGGTTCCTCATTTGGAAT 996

QY 563 TGGCCCTGTGACAGACTTCCCTTTTGTACACCCCTCTTGTGCTGTGTCTCATGTC 622
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 997 TGGCTGTGACAGACATCTCAATATGAGCTCATGTTATGTGACAGAGATGATATTC 1056

QY 623 TTTCTCTTCCCTTCTCATCATCATGAGCCCTCTATGCTTGCATCTCCTAGGGCTGTCTCC 682
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1057 TATGTCTCCCTCTCTCTGCAATTAATACCTCTTACTACTAATTTATTCACCTGTCTGC 1116

QY 683 GAATAGCTCTGTCTCAGGCGCTGAAAAAGCCCTGGCCACCTGCTCTCCACACCTAACAG 742
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1117 ACATGCTTCAATGAGGGAGAGAAAGCCCTGTGACCTGTTCCTCCACCTGTGACTG 1176

QY 743 CTGTACCCCTCTTCTATGAGGGGAGCCATGTTCATGTAAGGCTTACGGGCTACCGGG 802
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1177 TGTGTGGATGTCTATGAGGGGAGCCACTTTCATGTATGTGTGCGCCAGTTCCTTCCACA 1236

```

```

QY 803 CCCCTAGCCATGACAAGGTGGCTCTATCTTCTACACAGTCTCTACTCCCATGTGAACC 862
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1237 GTCTTACGACAGACAATATCATCTCTGTGTCTATACAAATGTTCACACAGCTGTGAACC 1296

QY 863 CCTCATTTACAGCTTGAAGATGGGAGCTGATGGGGCACTGAGGAAGGGGCTGG 919
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1297 CCTCATTTACAGCTTGAAGATTAAGAGGTGATTTGGACCTGTTAGAGAGTACTGG 1353

```

## RESULT 4

AK017036

LOCUS

DEFINITION

AK017036 1394 bp mRNA linear HTC 19-JAN-2002  
Mus musculus adult male testis cDNA, RIKEN full-length enriched  
library, clone:4933433E02:similar to T1 OLFATORY RECEPTOR, full  
insert sequence.

ACCESSION

AK017036.1 GI:12856091

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gliss, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
Quackenbush, J., Schraml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaudo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-Oka, K., Wang, K. H., Weltz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL

MEDLINE	21085660
PUBMED	11217851
REFERENCE	5 (bases 1 to 1394)
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carlinici, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schiml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:<http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)  
**COMMENT** Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cdNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cdNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTTVN 3'], cdNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cdNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cdNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAAATTATCCCCCCCCCCC 3']. cdNA was cleaved with BamHI and XhoI. cdNA of size comprised between 0.5 and 3 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES	Location/Qualifiers
source	1..1394
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:4933433E02"
	/db_xref="MGI:1894709"
	/db_xref="taxon:10090"
	/clone="4933433E02"
	/sex="male"
	/tissue_type="testis"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
CDS	332..1282

CDS

```
332 .1282
/note="data source:SPTR, source key:Q9EPF8, evidence:ISS
putative
similar to T1 OLFATORY RECEPTOR"
/codon_start=1
/protein_id="BAB30564.1"
/db_xref="GI:12856092"
/db_xref="MGI:1914036"
/translation="MEPNSTLESGFVLVGLDGSPELLCATVTTLYMLALISGDI
LLVITVDARLHVPMYLLRLSLIDLLFTSVPTNPNTVDFLRLONTISFEGCALDIE
SAMTLGGAEDLLAFMAYDRYVALCHPLNTYMIEMSPKACRLMAVIAISILASLALCHT
VYTMHEPFCSQEIIRHLICEVPLRLKACADDSOYELMAYVTGVFLLPLISAITSYV
SLILETVLHMPMSNEGKKALVPCSSHLTVAGMFGATFMVVLPSSEFHSFKQDNIISV
FTVITPALNPLIYSLRNKEVIGAVRRVLGRHILPAHATV"
```

BASE COUNT	355	a	337	c	304	g	398	t
ORIGIN								
Query Match	31.28; Score 295.4; DB 11; length 1394							

Best Local Similarity	58.1%;	Pred. No. 3e-65;
Matches	521; Conservative	0; Mismatches 376; Indels 0; Gaps 0;

QY	23	CCACACAGATGGCTTCTTCCCTCTTGGGCATCTTTCCACAGGCCAGACTGACCTTGTC	82
Db	351	CCTTGGAAAGTGGATTTCATCTTGGTGGGGATTCTGGATGGCAGTGGCTGCTCCTGAAGTGC	410
QY	83	TCTTCCTCGAGTTAAGTGGTCTTCACAGTGGGCCCTCTGTGGGAATGTCTCCTCATCT	142
Db	411	TCGTGGCCACAGTTACAAACCCTGTACATGTTGGCAGCTGATCAGCAATGGACTTCTACTCC	470
QY	143	TCTCATCTACCTGGACGCTGGACTTTCACACCCCATGTACTCTCTCCACGACGCTCT	202
Db	471	TGGTCATCACAGTGGATGCCCGGCTTCACGTAACCATGTACCTCTACTGAGGACGCTGT	530
QY	203	CCCTCATGGACCTCATGTGTGTCTGTAAACATGTGTGCCAAAGATGGAGCCCAACTTCCTGT	262
Db	531	CTTCATTTGACCTCTCTTCACATCAGTTGTAACTCCCAACACTGTGTGTGATTTCTGTC	590
QY	263	CTGGCAGGAGTTCATCTCTCTTTGTGGGCTGTGGCATACAAATTTGGCTTTTGTCTCTC	322
Db	591	TCAGAGACAAACACCATATCTCTTGAAGGATGTGCCCTTCAATTGTCTCAGCAATGACAT	650
QY	323	TTTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCTTATGACCACTACGTTGGCCGTTA	382
Db	651	TGGGTGGTGACAGAGACCTCTCTGTGGCCTTCATGGCCTATGATAGTATGTGGCCATTT	710
QY	383	GCCACCCACCTTCACTATCCCATCTCATGTAATCAGAGGGGTCTGTCCAGATTACTGGGA	442
Db	711	GTTCATCTCTTAACACTATGATCTTCATGAGTCCAAAGGCTCGACGCTCATGTGTGCCA	770
QY	443	GCTCCGCGGCTTTGGGATTAATAGATGGAGTGAATCAGATGTTGGCAGCCATGGGCTTAC	502
Db	771	TATCATGAGATCTTACGATCTCTTAGTGGCCCTAGTGCACACAGTGCACAAATGACACTTC	830
QY	503	CTTACTGTGGCTCAAGGAGCGTGGATCACTTTTCTGTGAGGTACAAAGCTTTATTGAAGC	562
Db	831	CTTTTTCATGTCTCCAGGAATTCAGACACCTGCTCTGTGAGGTCTCCATTTGTTAAAT	890
QY	563	TGGCCTGTGACAGACACTTCCCTTTTGACACCCCTCTTGTGCTGTGCTTCATGC	622
Db	891	TGGCTGTGACAGACATCTCAATATGAGCTCATGTTTATGTGACAGAGGTGATATTC	950
QY	623	TTTCTCTTCCCTTCTCCATCATCATAGGCCCTCCATGCTTGATCTCAAGGGGCTGTCTCC	682
Db	951	TATGTCTCCCCCTCTCTGCCATTTATACCTCTCTACTCACTAATTTCTATTCAGCTGTCTGC	1010
QY	683	GAATACGCTCTGCTCAGGCCCTGGAAGAACCCCTGGCCACCTGCTCTCCACCTTAACAG	742
Db	1011	ACATGCCCTTCAAAATGAGGGCAGGAGGAAGAACCCCTGTGCACCTGTCTCCACCTGACTG	1070
QY	743	CTGTACACCTCTTCTATATGGGGCAGCCATGTTCAATGTAACCTGAGGCCCTAGGGGCTTACC	802
Db	1071	TGGTTGGGATGTTCTATATGGGGGTGCCACTTTCATGTATGTGTGCTGCCAGTTCCTCCACA	1130
QY	803	CCCTTAGCCATGACAAGGTGGGCTCTATCTTCTACACAGTCCCTTACTCCCATGTGAACC	862
Db	1131	GTCTTAAGCAAGACAATATCATCTCTGTGTCTTAACAATTGTACACACCAGCTCTGAACC	1190
QY	863	CCCTCATTTACAGCTTGAGGAATGGGGAGGTGATGGGGGACACTGAGGAAGGGGCTGG	919
Db	1191	CCCTCATTTACAGCTTGAGGAATTAAGGAGGTGATGGAGCTGTGTAGAAGAGTACTGG	1247

RESULT 5	BC016940	LOCUS	BC016940	2021 bp	mRNA	linear	HTC 09-NOV-2001
DEFINITION	Homo sapiens, Similar to olfactory receptor, family 2, subfamily A, member 4, clone IMAGE:4424116, mRNA.						
ACCESSION	BC016940						
VERSION	BC016940.1 GI:16877381						
KEYWORDS	HTC.						
SOURCE	Homo sapiens.						







/clone="UUGC1M0562E19"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 124 a 195 c 129 g 184 t  
ORIGIN

Query Match 26.3%; Score 249.4; DB 17; Length 632;  
Best Local Similarity 63.9%; Pred. No. 1.5e-53;  
Matches 393; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

```
OY 166 CTTCACACCCCATGTACTCTCTCTCAGCCAGCTCTCCCTCATGACCCATGTTGTC 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 CTCACATACCAATGACTTTTATCAGCAGCTGTCCCTCATGACATGATGATACATT 78
OY 226 TGTACATTTGCCCAGAGAGGAGCCAACTTCTCTGTCGAGAGAGTCCATCTCTT 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 TCTGTCACTGTGCCCAAGATGCTCATGACCAAGTCTCTGGGAGCCACAAGATCTCAGCT 138
OY 286 GTGGGCTGTGGCATACAAATGGCTTTTCTCTCTCTGTGGGATCTGAGGGCTCTTG 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 GCTGCTGTGGGATGACAGATGCTCTCTACCTGACATTAGCAGGGTCAAGATATTTCTT 198
OY 346 CTGGGACTCATGGCTTATGACCACTACGTGGCCGTAGCCACCCTTCACTATCCATC 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 CTGCTGCCATGCTCTATGACCACTATGTGGCCATCTGTCTCATCCACTCCAGTATCTGTC 258
OY 406 CTCATGATCAGAGGCTCTCTCCAGATTACTGGAGCTCTGGCCCTTTGGGATAATA 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 CTCATGACCAACAGGGTGTCTCTCTCTGCTGATGTCTGTGTGGATCTCGGGATCTTGC 318
OY 466 GATGAGTGTATCAGATGTTGAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTC 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 GATGGCTCATGTTCACCCCTGTCAACATGACCTTCCATTTCTGTGATCCCGGAGATC 378
OY 526 GATCACTTTTCTGTGAGGTACAAGCTTTATGAAGCTGGCCCTGTGCAGACACTTCCCTT 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 CATCACTTCTCTGTGAGGTCCCTGTGTGACAAAGCTTCTCTGCTCAGACACTGGCTC 438
OY 586 TTGACACCCCTCTCTTGTGCTGTGTCTTCACTGCTTCTCTCTCTCTCTCTCTCTCTC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 TATGAGACCCCTCATGTATGTGTCTGTGTCTCATGCTTCTCTCTCTCTCTCTCTCTCT 498
OY 646 ATGGCCTCTATGCTTGCATCTTAGGGGCTGTGCTCCGAATAGCTCTGCTCAGGCTTG 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 TCAGGCTCTATACATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 558
OY 706 AAAAAAGCCCTGGCCACCTGCTCTCTCCCACTAAGCTGTCAACCTCTTCTATGGGCA 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 AAGAAGGCCCTTGCCACCTGCTCTCT-CCACATGACTGTGTCAACCTCTCTATGTGTCT 617
OY 766 GCCATGTTCAATGAC 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 618 GCTATCTATACCTAC 632
```

RESULT 7  
AG067024/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
SUBMITTED (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan (E-mail:chimps@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  
PRIMERS  
Sequencing: M13Rev  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI.  
Location/Qualifiers  
1. 650  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-056004.R"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 206 a 142 c 178 g 124 t  
ORIGIN

Query Match 23.4%; Score 221.4; DB 17; Length 650;  
Best Local Similarity 62.6%; Pred. No. 2.5e-46;  
Matches 345; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```
OY 143 TCCTCATACCTGAGCGCTGACTTCAACCCCATGTAATCTTCTCAGCCAGCTCT 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 TTCTTATCCAAACAGATTGCGCCCTTCAATACACCCATGTAATCTTCTCAGCCAGCTTT 549
OY 203 CCCTCATGAGCCTCATGTTGCTGTGAACATTTGCGCAAGATGGACCAACTTCTGT 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 CCTTAATGACATGATGATATTTTCCACTATTTGCTTAAGATGCTGTTAATTAACCTGC 489
OY 263 CTGGCAGAGATCCATCTCTTGTGGGCTGTGGCAATGAATTTGCTTTTGTCTCTC 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 TGGATCAAAAGGACCAATTTCTTGTGGGCTGACACAGCTCAACACTTCTTACCTTACC 429
OY 323 TTGTGGATCTGAGGGGCTCTGTGGGACTCATGCTTATGACCACTACGTGGCCCTTA 382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 TTGTGGAGCTGAATTTCTCTGTGGGCTCATGCTTATGACCGCTATGTGGCCATTT 369
OY 383 GCCACCACTTCACTATCCATCTCATGATGATGAGAGGCTCTGTCTCAGATTAATGGA 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 GCAACCTCTGAGATACCTGTCTCTGATGAGCCGCGGCTCTGTGATGATTAATGACAG 309
OY 443 GCTCTGGGCTTTGGGATATAGATGAGTGTATGATGATGATGATGATGATGATGATGAT 502
```

Db 308 GTTCCGTGGTTGGGGGCTTTTGGATGGCTCTCTAACCCTCACCATGAGCTTTC 249  
 QY 503 CTTACTGTGGCTCAAGAGCGGTGATCACTTTTCTGTGAGGTACAGCTTTATGAAGC 562  
 Db 248 CCTTCGCAATTCCTCCGGGAGATTAAACACTTCTCTGTGAGGCACAGCAGCTCTGAAGT 189  
 QY 563 TGGCCGTGAGACACATTCCTTTTGTAGACCCCTCTCTTGTGCTGTGCTTCATGC 622  
 Db 188 TGGCATGTGACAGACACAGCCCTCTATGAGACAGTGTATGTGTCTGTGTTTGTATGC 129  
 QY 623 TTCTCCTTCCTCTCTCCATCATCATGAGCCCTCTATGCTTGACATCCTAGGGGCTGTCTCC 682  
 Db 128 TGCTGATTCCTTCTCTGTAGTCTCTGCTTCTCTATGCCCGAATCCTGACTACAGTTCAGT 69  
 QY 683 GAATACGCTCT 693  
 Db 68 GCATGAGCTCT 58

RESULT 8  
 A0558499/c 592 bp DNA linear GSS 29-MAY-1999  
 LOCUS HS\_2082\_B1\_A05\_MR CIT Approved Human Genomic Sperm Library D Homo  
 DEFINITION sapiens genomic clone Plate=2082 Col=9 Row=B, DNA sequence.  
 ACCESSION A0558499  
 VERSION A0558499.1 GI:4918231  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 592)  
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 MEDLINE 99380589  
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).  
 BAC end Web Server: http://www.htsc.washington.edu  
 Plate: 2082 row: B column: 9  
 Seq primer: M13 Reverse  
 Class: BAC ends

FEATURES  
 source  
 1. 592  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_plate="2082 Col=9 Row=B"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /sex="male"  
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"  
 BASE COUNT 193 a 115 c 146 g 125 t 13 others  
 ORIGIN

Query Match 21.6%; Score 205; DB 17; Length 592;  
 Best Local Similarity 61.2%; Pred. No. 4.1e-42;  
 Matches 325; Conservative 0; Mismatches 206; Indels 0; Gaps 0;  
 QY 175 CCCATGTACTTCTCTCAGCAGCAGCTCTCCCTCATFGGAGCTCATGTGTGTATACATT 234  
 Db 592 CGCATGTACTTCTCTCAGTCACTGTCCTCCNCACTGTAGCTAGGCTCATCTGCACACG 533

QY 235 GTGCCAAGATGGCAGCCCACTTCCTGTCTGGCAGGAAGTCCATCTCTTGTGGCTCT 294  
 Db 532 GTACCCAGATGGCCCTTCAACTGCTGTCTGGCAGCAAGTCCATTTCTATGGCTGCTGT 473  
 QY 295 GGCATACAAATTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTCTTGGGACTC 354  
 Db 472 GCCACACAAATTATCTCTGTGTATACACTGCTTGGCTGTGAATGCTTCTGTGGCTGT 413  
 QY 355 ATGGCTTATGACACACATGAGCCGTGAGCCACCCACCTTCATATCCATCTCATGAAT 414  
 Db 412 ATGTCTTATGACTGCTCATATTCATTTGCCATTTGCCACCCCTGTAGATACACCAATCTCATGAGA 353  
 QY 415 CAGAGGCTGTCTCCAGATTACCTGAGAGCTCCCTGGGCTTTGGATATAGATGAGTGC 474  
 Db 352 CCCAANNTTGTAGACTTATGACTGCTCTCTCTGATCTTGGGCTTACAGATGANATC 293  
 QY 475 ATTCAGATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGATCACTTT 534  
 Db 292 ATTTATGCTGAAGCCACATTTTCTCTTCTACTGTGGGCTCTGGGAATAGCCCACTTC 233  
 QY 535 TTCTGTGAGTACAACTTTATTTGAAGCTGGCCTGTGCAGACACTTCCCTTTTGACACC 594  
 Db 232 TTCTGTGAGTAACTTCCCACTAATCTCTCATGCAATGACACATCAATATCTGAAAG 173  
 QY 595 CTCCTCTTGGCTGTCTGTCTCATGCTTCTCTCCCTTCTCCATCATCATGAGCCCTCC 654  
 Db 172 GTTATTTTCATGTGCTCTATAGTAATGCTTGTCTTCCCTGTGCAATCATCATTTGCTTCA 113  
 QY 655 TATGCTTGCAATCTAGGGGCTGTGCTCCGATACGCTCTGCTCAGGCTGG 705  
 Db 112 TATGACAGAGTAATTCTGGCTGTCTATTCACATGGNATCTGAGAGAGGCTCGG 62

RESULT 9  
 A0425578/c 448 bp DNA linear GSS 24-MAR-1999  
 LOCUS CITBI-E1-2567E19.TF CITBI-E1 Homo sapiens genomic clone 2567E19,  
 DEFINITION DNA sequence.  
 ACCESSION A0425578  
 VERSION A0425578.1 GI:4495666  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 448)  
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.  
 TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: CITBI-E1-2567E19.TR  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.

FEATURES  
 source  
 1. 448  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2567E19"  
 /clone\_lib="CITBI-E1"  
 /sex="male"  
 /cell\_type="sperm"  
 /note="Vector: pBelobAC11; Site\_1: EcoRI; Site\_2: EcoRI;"

Query Match 21.6%; Score 205; DB 17; Length 592;  
 Best Local Similarity 61.2%; Pred. No. 4.1e-42;  
 Matches 325; Conservative 0; Mismatches 206; Indels 0; Gaps 0;  
 QY 175 CCCATGTACTTCTCTCAGCAGCAGCTCTCCCTCATFGGAGCTCATGTGTGTATACATT 234  
 Db 592 CGCATGTACTTCTCTCAGTCACTGTCCTCCNCACTGTAGCTAGGCTCATCTGCACACG 533

Caltech Human BAC Library D"  
BASE COUNT 149 a 80 c 126 g 93 t  
ORIGIN

Query Match 21.2%; Score 200.6; DB 17; Length 448;  
Best Local Similarity 66.1%; Pred. No. 4.9e-41;  
Matches 290; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

```
QY 104 TCTTCACAGTGGCCCTCTGTGGGAATGTCCCTCATCTTCCTCATCTACCTGGACGCTG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 TCTTTTAGTGCATTCAGGGAAGTCTGTGCATGTCTTCCTCATATACCTGGACACCC 385
QY 164 GACTTCACACCCCATGTACTTCTCTCCAGCCAGCTCTCCCTCATGACCTCATGTGG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 AGCTCCACACCCCATGTACTTCTCTCCAGCCAGCTCTCCCTCATGACCTCATGTGCA 325
QY 224 TCTGTAAATGTGCCAAGATGGCAGCACTTCTGTGGCAGGAATCCATCTCCT 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 TCTGCACACCCGTACCCAGATGGCCTTCACTACCTGTCTGGCAGCAAGTCCATTTCTA 265
QY 284 TTGTGGCTGTGGCATACAAATTTGGCTTTTGTCTCTCTGTGGGATGTGAGGGGCTCT 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 TGGCTGTGTGTGCACACAATTTTCTCTATATATACATGTCTGGCTGTGATGTTC 205
QY 344 TGCTGGACATGCTTATGACCACTAGCGGCCGTAGCCACCCACTTCATATCCCA 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 TTTTGGCTGTATGCTTATGACCGCTATATGCTATTTGCCACCCCTTAAGATATACCA 145
QY 404 TCTCATGAATCAGAGGCTGTCTCCAGATTAATGAGGCTCTGGGCTTTGGGATAA 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 ATCTCATGAATCCTAAATTTGTGACTTATGGCTACCTTCTCTGTGATCCTGGGCTCTA 85
QY 464 TAGATGAGTGAATTCAGATGCTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCG 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 CAGATGGAATCATGTATGCTGTAGCCACAAATTTTCTCTTCTTGTGGGCTCGGGA 25
QY 524 TGGATCACTTTTCTGTGA 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24 TAGCCACTTCTTCTGTGA 6
```

RESULT 10  
B0888002 915 bp mRNA linear EST 16-AUG-2002  
LOCUS  
DEFINITION AGENCOURT\_8742018 NIH\_MGC\_129 Mus musculus cDNA clone IMAGE:6390379  
5', mRNA sequence.

ACCESSION B0888002  
VERSION B0888002.1 GI:22280016  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 915)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabp-remail.nih.gov  
Tissue Procurement: Susan L. Sullivan, PhD.  
CDNA Library Preparation: Resgen, Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13876 row: 0 column: 20  
High quality sequence stop: 700.

FEATURES  
source  
1. 915  
Location/Qualifiers

/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6390379"

/clone\_11b="NIH\_MGC\_129"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: olfactory epithelium; Vector:  
pCMV-Sport6.1.ccd; Site\_1: EcoRV; Site\_2: NotI; Cloned  
unidirectionally. Primer: Oligo dT. Average insert size  
2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this  
is a NIH\_MGC Library."

BASE COUNT 181 a 287 c 201 g 246 t  
ORIGIN

Query Match 20.9%; Score 198; DB 14; Length 915;  
Best Local Similarity 53.7%; Pred. No. 3.3e-40;  
Matches 433; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

```
QY 124 GGGATGTCCCTCCTCATCTTCTCATCTTACCTGGACGCTGACCTTCACACCCCATGTAC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 GGAACCTGCTCATTTATGGTACATCATGACGACGACCCCAACCCCTGCATACCCCTATGTAC 63
QY 184 TTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTGTGCTGTACATTTGTGCCAAG 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 TTCTTCTGTCACACCTGCTGTGTGTGTGACATCTTGTGACCTCCACCATCTACCCAG 123
QY 244 ATGGCAGCACTCTCTGTGGCAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAA 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 CTACTGACACCATGTGTGGGGGAAGA---CCATCTTTACGGGGGCTGATGGCCAG 180
QY 304 ATTGGCTTTTGTCTCTCTGTGGATCTGAGGGGCTCTGCTGGACATGAGCTTAT 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CTCTTCTTCTTACATGGTACATGGGGGTGAGAGCTTCTCTCTCAGCTATGGGCTAT 240
QY 364 GACCACATGAGTGGCGGTAGCCACCCACTTTCATATCCATCTCATATGATCAGAGGCTC 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GACCGCTTTGTGGCCATCTGCTGCCCCCTGTGATATATAGTACCTGGATGGGCCCCAGAGTG 300
QY 424 TGTCTCCAGATTTACTGGGAGCTCTGGGCTTTGGGATATAGATGAGATTCAGATG 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TGTGATTTCTGGCTGGCAATTTGCTGTGGCTATCAGCCCTGACTAACACCAAGCATAACTGA 360
QY 484 GTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCTGATCACTTTTCTGTGAG 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AGCCTGTGCTACGTCTACATTTCTGACAGCTCCAAATGTGTTGAGCACTTCTTCTGTGAG 420
QY 544 GTACAAGCTTTATGAAGCTGGCTGTGAGACACACTTCCCTTTTGTACACCCCTCTCTTT 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 ATTCCCCCACTGTGAAGCTCTCTTGTGCTCCCAACACAAATGAACGAGGCTATGCTTTC 480
QY 604 GCTTGTGTCTTCTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 GCTGACAGATGTGTCTCTGCTGTAGGGAATCTCTGTGATCATCTCTCTCTATAGGCTTT 540
QY 664 ATCTAGGGGCTGTGCTCGAATACGCTGTCTGAGGCTTGAAAAAGCCCTGGCCACC 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 ATTGTGGCAGCATCTCTGAAGATCCGCTGAGCTGAGGGGCAAGCAAGAGCTTCTCTCAC 600
QY 724 TGCTCTCCCACTTAACAGCTGTACACCTCTTCTATGGGGCAGCCATGTTTCATGTACTG 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 TGCTCTGACACCTCATGCTGTGTCACCATGTACTACTCTGATCATCTGATACCTACATTT 660
QY 784 AGGCTAGGCGCTACCGGGCCCTAGCCATGACAAAGGTGGCTCTATCTTCTACACAGTC 843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 CGCCCTTCACTCAGCTTACTCTGACCAAGGAGGAGGTGTCATCTTCTACACCTCA 720
QY 844 CTTACTCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTATGGGGCA 903
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 GTGGACCCACCCCTTGACCTCTCATCTTACACTGTGAGGAATAGGATGTCCAAAGTTGCA 780
QY 904 CTGAGGAAGGGGCTGAGCGGCTGACG 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 CTCCGAGAGCTTTTCTCTGCTGTGCTG 806
```

RESULT 11  
BI912805

LOCUS B1912805 518 bp mRNA linear EST 16-OCT-2001  
 DEFINITION 603176272F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5240705 5',  
 mRNA sequence.  
 ACCESSION B1912805  
 VERSION B1912805.1 GI:16177058  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 518)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM11606 row: p column: 18  
 High quality sequence stop: 516.  
 Location/Qualifiers  
 source 1..518  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5240705"  
 /clone\_1id="NIH\_MGC\_121"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3  
 fetal brains, female age 20 weeks, female age 24 weeks,  
 and male age 26 weeks. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 0.7-3.5 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 017. Note:  
 this is a NIH\_MGC Library."  
 BASE COUNT 99 a 155 c 103 g 161 t  
 ORIGIN

Query Match 20.6%; Score 195; DB 13; Length 518;  
 Best Local Similarity 61.5%; Pred. No. 1.5e-39;  
 Matches 312; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

OY 44 TCTTGGGCACTTTTCCACAGCCAGACTGACCTTGTCCTCTCTGACGTTATGGTGG 103  
 ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |  
 Db 12 TGTGGGTGCTCTCCCAATCAAACTGGAATATTTCTTGTGCTTATCATCTCTCA 71  
 OY 104 TCTTCACAGTGGCCCTCTGTGGGAATGTCTCTCATCTTCTCATCTTACCTGAGCGTG 163  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 72 TATTTCTTGTGGCTCTGGTGGGTAACCTGGCCATGATTCACCTCATCCACGTGATCCTC 131  
 OY 164 GACTTCACACCCCATGTACTTCTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGG 223  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 132 GTCTCCACACACCGATGTACTTCTTCTCAGCCAGCTCTCCCTTATGAGACCTGATGACA 191  
 OY 224 TCTGTAACTGTGCCAAAGATGGCAGCACTTCTGTGTGGCAGGAAGTCCATCTCTCT 283  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 192 TCTCCACACACCGTCCCAAGATGGCGTACAACTTCTGTCCGGCCAGAAAGGCACTCTCT 251  
 OY 284 TTGTGGGCTGTGGCATACAATTTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTCT 343  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 252 TCCTGGGATGTGTGCAAGCTTCTTCTTCTGACCATGGCGTGTCTGAAGGCTTAC 311  
 OY 344 TGCTGGAGCTGATGGCTTATGACCACTAGCGGCGGTAGCCACCCACTTCACTATCCCA 403  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 312 TCCTGACCTCATGCGCTACAGACCGTTATTTGGCCATCTGCCCACTCTCTATTTATCCCTA 371

OY 404 TCCTCATGAAATCAGAGGCTGTCTCTCCAGATTACTGGAGCTCTGGGCTTTGGATAA 463  
 ||||| ||||| | | | | | | | | | | | | | | | | | | | | | |  
 Db 372 TCCGCATGAGTAATAATGATGTGTGAGAGATGATTGGAGCGCTTTGACACATGGGGTCCA 431  
 OY 464 TAGATGAGATTTAGATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCG 523  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 432 TCAACTCCTTGGCACACACAGCTTTTGGCCCTTCATATTTCCCTACTGCAGAGTCTAGGGCTA 491  
 OY 524 TGCATCATTTTTCTGTGAGGTACAAG 550  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 492 TTGACCATTCTTCTGCGATGCCAG 518

RESULT 12  
 B0892087  
 LOCUS B0892087 1046 bp mRNA linear EST 16-AUG-2002  
 DEFINITION AGENCOURT\_8743862 NIH\_MGC\_129 Mus musculus cDNA clone IMAGE:6389703  
 5', mRNA sequence.  
 ACCESSION B0892087  
 VERSION B0892087.1 GI:22284101  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1046)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Susan L. Sullivan, PhD.  
 cDNA Library Preparation: Resgen, Invitrogen Corp.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13875 row: c column: 16  
 High quality sequence stop: 559.  
 Location/Qualifiers  
 source 1..1046  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6389703"  
 /clone\_1id="NIH\_MGC\_129"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: olfactory epithelium; Vector:  
 pCMV-SPORT6.1.cdb; Site\_1: EcoRV; Site\_2: NotI; Cloned  
 unidirectionally. Primer: Oligo dT. Average insert size  
 2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this  
 is a NIH\_MGC Library."  
 BASE COUNT 222 a 300 c 235 g 289 t  
 ORIGIN

Query Match 19.8%; Score 187.4; DB 14; Length 1046;  
 Best Local Similarity 54.8%; Pred. No. 2e-37;  
 Matches 371; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

OY 164 GACTTCACACCCCATGTACTTCTTCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGG 223  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 1 GGCTGCACACACCAATGATTTCTTTGTGGCAACCTCTCATTTCTGACATGTGTTTA 60  
 OY 224 TCTGTAACTGTGCCAAAGATGGCAGCACTTCTGTGTGGCAGGAAGTCCATCTCTCT 283  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 61 CCACGGTTTCCATCCACAGATGTCTTACAACTCTGGGCTCCAGACAAAGACATCAGCT 120  
 OY 284 TTGTGGGCTGTGGCATACAATTTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTCT 343  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 121 ATGTGGTGTGTATCAACTGTACTTGTCTTAGCCCTGGGAGGGGTAGATAATGTCTC 180  
 OY 344 TGCTGGAGCTCATGGCTTATGACCACTAGCGTGGCGGTAGCCACCCACTTCACTATCCCA 403  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

```
Db 181 TCCTGGCTGTATGACATATGACCGCTAGCGCGCTGTCTGCAAGCCCCCTGCACACTACACCG 240
QY 404 TCCTCATGAATCAGAGGGTCTGTCTCCAGATTACTGGAGCTCCTGGCCCTTGGGATAA 463
Db 241 TCATCATGATCTCTGCTCTGCGGACAGCTGCGCTTCAATGCGATGTTGAGCGGCTTGG 300
QY 464 TAGATGAGTATGATTCAGATGGTGGCAGCCATGGGCTTACTTACTGTGGCTCAAGAGCG 523
Db 301 GCAATTCCTCATCATGACGACCCAGACGCTGATGTGCCCCGCTGTGACACAGAGCGGG 360
QY 524 TGGATCATTCTTCTGTGAGGTACAAAGCTTATTTGAAGCTGGCCTGTGACAGACACTTCCC 583
Db 361 TGGACCACTTCTCTGTGACATGCGCGCCCTGATTTGGCATGGCCCTGTGTGACACTATGG 420
QY 584 TTTTGAACCCCTCCTCTTGTGCTGTCTGTCTCATGCTTCTCTCCTTCCCTTCTCATCA 643
Db 421 CCTTGAAGACACTGGCTTGTGCGTGGCGAATTTTATCATCTTGGACCGCTCATCTCA 480
QY 644 TCATGGCCCTCTATGCTTGCATTCCTAGGGGCTGTGCTCCGAATACGCTGTGCTCAGGCT 703
Db 481 TCCTTATCTCCTATGTTACATTTGCCGACAGATGTTAGATCAAGTCAAGCTGTGCGG 540
QY 704 GGAAGAAAGCCCTGGCCACCTGCTCCACCTAACAGCTGTACCCCTCTTCTATGGGG 763
Db 541 GACGGAAGCCCTCAACACACTTGACGCTCCACCTTATCGTTGTCTCTCTCTTACGGTA 600
QY 764 CAGCCATGTTATGATACCTGAGGCTTAGGGCTACCGGCGCTAGCCATGACAGGTGG 823
Db 601 CCATCATATATACATGATACCTCCAGCGCGCAACCTTACTCCCGACAGAGGGGCAAGTT 660
QY 824 CCTCTATCTTCTACACA 840
Db 661 CTCACCTCCTTCTTCTACA 677

RESULT 13
LOCUS AZ868627 618 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0180J22F Mouse 10kb plasmid UGCM library Mus musculus genomic
c1one UGCC2M0180J22 F, DNA sequence.
ACCESSION AZ868627
VERSION AZ868627.1 GI:13072131
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 618)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0180 row: J column: 22
Seq primer: CGTTGTAAGAGAGAGAGAGAGAGT
Class: plasmid ends
High quality sequence stop: 618.
FEATURES
source 1..618
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
```

```
/clone="UGCC2M0180J22"
/clone_lib="Mouse 10kb plasmid UGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1473214|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

```
BASE COUNT 116 a 189 c 139 g 174 t
ORIGIN
Query Match 19.5%; Score 184.6; DB 17; Length 618;
Best Local Similarity 59.1%; Pred. No. 7.8e-37;
Matches 316; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 56 TTTCCACAGCCAGACTGACCTGTCTCTCTCTCTGACATATAGTGCTTCACAGTG 115
Db 73 TTAGCTACACAGGGCCCCCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 132
QY 116 CCCTCTGTGGGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 175
Db 133 GCCTGTAGGCAACACACACCCCTGCTTTTCTGTGATTTGCCACAGACTCCAGCTC 192
QY 176 CCATGTACTTCTCTCTACGCCAGCTCTCCCTCATGAGCCTCATGTTGGTCTGTA 235
Db 193 CCATGTACTTCTCTCTACGCCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252
QY 236 TGCCAAAGATGGCAGCCCACTTCTGTCTGCGCAGGAAGTCCATCTCTTTGGCT 295
Db 253 TCCCAAGTGGTGGCTGAGTCTCTGCAAGGACAGAAATGATATAGCTTTGGGGAT 312
QY 296 GCATACAAATTGGCTTTTGTCTCTCTGTGTGGATCTGAGGGGCTTCTGAGCTCA 355
Db 313 CCACTCAGATGTTCTTCTCTATGCTCATGGGTGTCTGAGGGGTCTCTCTCTCT 372
QY 356 TGGCTATGACCACTACGTGGCGGTAGCGACCCACCTTCACTATCCCATCTCATGA 415
Db 373 TGTCTTATGACCGCTACGTGGCGGTATGCGACCCCTGCACTACAGAGTGTCTAT 432
QY 416 AGAGGCTGTCTCTCAGATTAATCTGAGAGCTCTGGGCTTTGGATTAATAGATGA 475
Db 433 ACCAAGTGTGTCTGTGATGTTGGGTGATCTCTGTCTCAAGGTGCTTGGTGGCT 492
QY 476 TTCAGATGTTGGCAGCCATGGGCTTACTTCTGTGCTCAAGAGAGCGTGGATCATT 535
Db 493 TCCTGACCTCCATCACCCTGAGTGTCCCTTACTGTGCTCACACACCGTGGACCAT 552
QY 536 TCTGTGAGGTACAAAGCTTATTTGAAGCTGGCGTGTGACAGACACTTCCCTTTTGA 590
Db 553 TCTGTGATGCGCAGCTCTCTCAAGCTGTCTCATGTGACAGACACATCAGCCTATGA 607

RESULT 14
LOCUS BG197640 796 bp mRNA linear EST 21-APR-2001
DEFINITION RST17016 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG197640
VERSION BG197640.1 GI:13719455
```



KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 796)  
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Kluka, A., Hess, J., Cochren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.  
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 544.  
Location/Qualifiers  
1. 796  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."  
BASE COUNT 115 a 271 c 189 g 220 t 1 others  
ORIGIN

Query Match 19.3%; Score 183.2; DB 12; Length 796;  
Best Local Similarity 53.3%; Pred. No. 2.1e-36;  
Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

QY 11 GGGTGACCACTCTACACAGATGGCTTCCTCTTGCGCATCTTCCACAGCCAGA 70  
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 20 GGGTAACCCACCTCCATGCTGATTCATCTCGTCTCTCTGCTTCCCCACC 79  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 71 CTGACCTGTCTCTCTCTGCAAGTATGCTGCTTCACAGTGGCCCTCTGGGAATG 130  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 80 TCCAACTGATGCTCTCTCTGCTGCTGCTGATGATGATGATGATGATGATGATG 139  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 131 TCCTCTCATCTTCTCTCATCTACCTGGACGCTGGACTTCACACCCCATGTAATCTTCC 190  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 140 TGTTCATCATGGCCACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 199  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 191 TCAGCCAGCTCTCCCTCATGAGACCTCATGTTGCTGTGAATGTTGCCAAGATGGCAG 250  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 200 TGTGCGTCTCTCAGTCTCGAGATCCTCTACACCGTGCCCATCATCCCGCATGCTGG 259  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 251 CCACTTCTCTGCTGGAGAGAGTCCATCTTGTGGGCTGTGGCATACAAATTGGCT 310  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 260 CCGACCTGCTGCTCCACCCAGCGCTCCATCGCTTCTGCGCTGTGCCAGTGAATGTTCT 319  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 311 TTTTGTCTCTCTGTTGGGATCTGAGGGGCTCTGCTGGGACTCATGGCTTAATGACCACT 370  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 320 TCTCTTCAAGCTTGGGCTTCAACCCACTCTCTCTGCTCACCCTCATGGGCTACGACCGCT 379  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 371 ACGTGGCGGTAGCCACCCACTTCACTATCCATCTCTCATGATCAGAGGGCTCTCTCC 430  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 380 ACGTGGCATCTGCGACCCCTGCGCTACACGAGCTCATGAGCCACGCGGCTGCGCT 439  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 431 AGATTACTGGAGCTCCTGGGCTTGGGATATAGATGAGTGAATTCAGATGCTGGCAG 490  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 440 GCCTGTGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 499  
111 111 111 111 111 111 111 111 111 111 111 111 111 111

QY 491 CCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGATCACTTTTCTGTGAGTACAG 550  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 500 TTTTCAACTGACTTCTGTGTGATTCAGAGATCCAGCATTTTATGTCATGTGCCAC 559  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 551 CTTTATGAGCTGGCGCTGTGACAGACACTTCCCTTTTTCACACCCCTCTTGTGTC- 609  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 560 CTCTGTGAAGTTGGCGCTGTGGAATAATGTACACAGCTGTGGCCCTGNCCTGGGCTTG 619  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 610 --TGTGTTTCATGCTTCT 667  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 620 TATGATCATGGCAGCTGTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 679  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 668 TAGGGCTGTGCTCCGATACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 727  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 680 TGCCGACATCTTGAAGATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 739  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
QY 728 CTTCCCACTACAGCTGTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 772  
111 111 111 111 111 111 111 111 111 111 111 111 111 111  
Db 740 CTTCTACCTTAATTGGGCTCATTTGTGACACTATGCGCTTCTCTCTCTCTCTCTCTCT 784  
111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 15  
AQ970130/c 481 bp DNA linear GSS 28-JAN-2000  
LOCUS AQ970130/c  
DEFINITION RPCI-23-314F11.TU RPCI-23 Mus musculus genomic clone RPCI-23-314F11  
ACCESSION AQ970130  
VERSION AQ970130.1 GI:6800503  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 481)  
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-23  
JOURNAL Unpublished (1999)  
COMMENT Other\_GSSs: RPCI-23-314F11.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
Plate: 314 row: F column: 11  
Seq primer: SP6  
Class: BAC ends.  
FEATURES  
source Location/Qualifiers  
1. 481  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-314F11"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
BASE COUNT 129 a 118 c 133 g 100 t 1 others





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2003, 03:35:05 ; Search time 73 Seconds  
(without alignments)  
3982.597 Million cell updates/sec

Title: US-09-634-109D-1

Perfect score: 948  
Sequence: 1 atgggaagatgggtgaacca.....ggatgtgcagccagcactga 948

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	388.8	41.0	1290	2	US-08-827-291A-1	Sequence 1, Appli
2	245	25.8	1713	2	US-08-467-948A-1	Sequence 1, Appli
3	245	25.8	1713	3	US-08-467-947A-1	Sequence 1, Appli
4	235.2	24.8	966	3	US-08-748-506-7	Sequence 7, Appli
5	233.6	24.6	966	3	US-08-748-506-8	Sequence 8, Appli
6	233.2	24.6	966	3	US-08-748-506-5	Sequence 5, Appli
7	227.4	24.0	966	3	US-08-748-506-6	Sequence 6, Appli
8	208	21.9	1080	4	US-09-668-680-9	Sequence 9, Appli
9	205.4	21.7	900	4	US-09-085-371-5	Sequence 5, Appli
10	203.4	21.5	1297	4	US-09-668-680-11	Sequence 11, Appli
11	201.4	21.2	1062	4	US-09-668-680-10	Sequence 10, Appli
12	186.2	19.6	984	3	US-08-748-506-9	Sequence 9, Appli
13	122.4	12.9	963	4	US-09-605-785-526	Sequence 526, App
14	122.4	12.9	963	4	US-09-439-313-526	Sequence 526, App
15	122.4	12.9	1474	1	US-08-465-980-1	Sequence 1, Appli
16	122.4	12.9	1474	2	US-09-053-303-1	Sequence 1, Appli
17	122.4	12.9	1474	5	US-09-339-115-1	Sequence 1, Appli
18	122.4	12.9	1474	5	PCT-US95-07093-1	Sequence 1, Appli
19	112.8	11.9	1828	3	US-08-988-876-2	Sequence 2, Appli
20	105	11.1	1539	4	US-09-668-680-13	Sequence 13, Appli
21	75.2	7.9	1320	1	US-08-599-252-84	Sequence 84, Appli
22	75.2	7.9	1320	1	US-08-332-074-57	Sequence 57, Appli
23	75.2	7.9	1320	5	PCT-US96-06352-84	Sequence 84, Appli
24	75.2	7.9	1320	5	PCT-US96-06583-84	Sequence 84, Appli
25	67	7.1	897	4	US-09-668-680-5	Sequence 5, Appli
26	62.6	6.6	7218	1	US-08-232-463-14	Sequence 14, Appli
27	56.8	6.0	3211	2	US-08-574-959A-8	Sequence 8, Appli

C	28	56.8	6.0	3211	4	US-09-357-014-8	Sequence 8, Appli
C	29	56.8	6.0	3901	2	US-08-574-959A-6	Sequence 6, Appli
C	30	56.8	6.0	3901	4	US-09-357-014-6	Sequence 6, Appli
C	31	47.8	5.0	251	4	US-09-605-785-352	Sequence 352, App
C	32	47.8	5.0	251	4	US-09-439-313-352	Sequence 352, App
C	33	47.8	5.0	251	4	US-09-352-616A-352	Sequence 17, Appli
C	34	47.6	5.0	1270	4	US-08-387-805-1	Sequence 1, Appli
C	35	46.6	4.9	921	1	US-08-722-001-17	Sequence 24, Appli
C	36	46.6	4.9	1567	1	US-08-722-001-24	Sequence 3, Appli
C	37	46.6	4.9	1738	1	US-08-334-698-3	Sequence 3, Appli
C	38	46.6	4.9	1738	1	US-08-228-932-3	Sequence 3, Appli
C	39	46.6	4.9	1738	1	US-08-468-939-3	Sequence 3, Appli
C	40	46.6	4.9	1738	2	US-08-406-855A-3	Sequence 3, Appli
C	41	46.6	4.9	1738	2	US-08-722-190-3	Sequence 3, Appli
C	42	46.6	4.9	1738	3	US-08-244-354-3	Sequence 3, Appli
C	43	46.6	4.9	1738	3	US-09-206-899-3	Sequence 3, Appli
C	44	46.6	4.9	1738	4	US-09-444-783-3	Sequence 3, Appli
C	45	46.6	4.9	1738	4	US-09-688-415-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-08-827-291A-1  
; Sequence 1, Application US/08827291A  
; Patent No. 5874243  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Sathe, Ganesh  
; TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: PA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/827, 291A  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1290 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-827-291A-1  
Query Match 41.0%; Score 388.8; DB 2; Length 1290;  
Best Local Similarity 64.4%; Pred. No. 7.8e-103;  
Matches 582; Conservative 0; Mismatches 322; Indels 0; Gaps 0;  
OY 9 ATGGTGAACCACTCTACACAGATGGCTTCTCTCTGGGCATCTTTCCACACGCA 68

Db	301	AAGGAGGAATTGACCTTCAACCTCCGACTTCATCTTCTGGGAATCTTCAATACACAGCCC	360
Qy	69	GACTGACCTTGTCTCTCTCTCGACAGTATAGTGTCTTACAGTGGCCCTCTGTGGAA	128
Db	361	CACCCACACCTTCTCTCTTCTGTGTCTGTGGCAATCTTTTACGTGGCCTTCATAGGGAAA	420
Qy	129	TGTCTCTCATCTTCTCTCATCTACCTGGACGCTGGAAGCTTACACCCCATGTACTCTT	188
Db	421	CTCTGTATGTTCTCTCTCATCTACCTGGACACCCAGCTCCACACCCCATGTACTCTCT	480
Qy	189	CCTCAGCCAGCTCTCCCTCATGAGCCTCATGTGTGTGTGAACAATGTGCCAAGATGGC	248
Db	481	CCTCAGCCAACTGTCCCTCATGAGCCTCATGTGTGTGTGAACAATGTGCCAAGATGGC	540
Qy	249	AGCCAACTCTCTGTGTGGCAGGAAGTCATCTCTTGTGTGGCTGTGGCATACAATTGG	308
Db	541	CTTCAACTACCTGTGTGGCAGGAAGTCATCTCTTGTGTGGCTGTGGCATACAATTGT	600
Qy	309	CTTTTGTCTCTTGTGTGGATGAGGGGCTCTGTGGCCTCATGTGCTTAAGACCA	368
Db	601	CTTCTATACATCACTGTCTGTGGATGAGGGGCTCTGTGGCCTCATGTGCTTAAGACCG	660
Qy	369	CTAGTGGCCGTTAGCCACCCACTTTCATCTATCCCATCTCATGATCAGAGGCTGTCT	428
Db	661	CTACACTGCCATTTGGCCACCCTCTAAGATACACCAATCTCATGAGCCCTAAATTTGTG	720
Qy	429	CCAGATTACTGGGAGCTCTGTGGCCTTTGGGATTAATAGATGAGTGTATCAGATGGTGGC	488
Db	721	ACTTATGACTGCCCTTTCTGTGATCTGTGGCTCTACAGATGGAATCATTTATGTCTAGC	780
Qy	489	AGCCATGGGCTTACCTTACTGTGGCTCAAGAGGCTGGATCACTTTTCTGTGAGGTACA	548
Db	781	CACATTTTCTTCTTCTTCTACTGTGGGCTCTGGGAATAGCCCACTTCTTCTGTAGTTACC	840
Qy	549	AGCTTATTTGAAGCTGGCCTGTGGAGACACACTTCCCTTTTGAACCCCTCTTGTGCTTG	608
Db	841	TTCCCTACTAATCTCTCATGCAATGACACATCAATATTTGAAGAAGTTATTTTCATTTG	900
Qy	609	CTGTGTCTCATGCTTCTCTCCCTTCTCCATCATCATGAGGCTTCTATGCTTGATCTCT	668
Db	901	CTGTATAGTAATGCTTGTCTTCCCTGTGTGCAATCATCATTTGCTTCTATGCTGAGTTAT	960
Qy	669	AGGGGCTGTGCTCCGAATAGGCTGTGCTAGGCCCTGGAAAAAGCCCTGGCCACTGCTC	728
Db	961	TCTGGCTGCATTCATGAGATGAGATCTGGAGAGGGTGTCCGAAGCTTTCACGACCTGTT	1020
Qy	729	CTCCACCTAACAGCTGTCAACCTCTTCTATGGGGCAGCCATTTTCATGTACTGAGGCC	788
Db	1021	CTCTACCTCATGTGTGGGTAATGTTCTATGAGCAGGTTTGTTCATGTACTGATACAGCC	1080
Qy	789	TAGGCGCTACGGGCCCCCTAGACATGACAAGTGGGCTCTATCTTCTACACAGTCTTAC	848
Db	1081	CACATCTGATCGCTTCCCAAGCAGGACAAGCTGTGTCTGTATCTACACCAATCCTCAC	1140
Qy	849	TCCCATGCTGAACCCCTCATTTTACAGCTTGAGAAATGGAGGTGATGGGGCACTGAG	908
Db	1141	TCCCATGCTGAATCCCTCATCTTACAGCCTCCGCAACAAGGAGTGCACAGACATTCAT	1200
Qy	909	GAAG 912	
Db	1201	GAAG 1204	

RESULT 2  
US-08-467-948A-1  
; Sequence 1, Application US/08467948A  
; Patent No. 5998164  
; GENERAL INFORMATION:  
; APPLICANT: LI, YI  
; APPLICANT: CAO, LIANG  
; APPLICANT: NI, JIAN  
; APPLICANT: GENTZ, REINER

```

1 APPLICANT: BULT, CAROL J.
2 APPLICANT: SUTTON III, GRANGER G.
3 APPLICANT: ROSEN, CRAIG A.
4 TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
5 TITLE OF INVENTION: Coupled Receptor GPR2
6 NUMBER OF SEQUENCES: 30
7 CORRESPONDENCE ADDRESSES:
8 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
9 STREET: 1100 NEW YORK AVE., NW, SUITE 600
10 CITY: WASHINGTON
11 STATE: DC
12 COUNTRY: USA
13 ZIP: 20005
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: FLOPPY DISK
16 COMPUTER: IBM PC COMPATIBLE
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/467, 948A
21 FILING DATE: 06-JUN-1995
22 CLASSIFICATION: 435
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: PCT/US95/04079
25 FILING DATE: 30-MAR-1995
26 ATTORNEY/AGENT INFORMATION:
27 NAME: STEFFEE, ERIC R.
28 REGISTRATION NUMBER: 36,688
29 REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 202-371-2600
32 TELEFAX: 202-371-2540
33 INFORMATION FOR SEQ ID NO: 1:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 1713 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: both
38 TOPOLOGY: both
39 MOLECULE TYPE: cDNA
40 FEATURE:
41 NAME/KEY: CDS
42 LOCATION: 116..1003
43 UN-08-467-948A-1
44
45 Query Match          25.8%; Score 245; DB 2; Length 1713;
46 Best Local Similarity 56.8%; Pred. No. 3.2e-61;
47 Matches 489; Conservative 1; Mismatches 366; Indels 5; Gaps 2;
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

Db 518 TATTTCATCATCATGACCTGGAAAAGTTCATCATCTCTGGGCATCACCTTCCCTGGACATGT 577

QY 457 GGGATATATAGATGAGCTGATTCAGATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCA 516

Db 578 GGCTCCCTCCTGGCTATGGTCCATGTGAGCCCTCATCTCAAGACTGCCCTTTTGTGGGCT 637

QY 517 AGGAGCGTGGATCACCTTTTCTGTGAGGTACAAAGCTTATTTGAAGCTGGCCCTGTGCAGAC 576

Db 638 CGTGAATCAACCACTTCTCTGTGAATCCTGTGTCTCCTCAGGCTGGCCTGTGCTGAT 697

QY 577 ACTTCCCTTTTGTACACACCCTCCTCTTGTCTGCTGTGTCTTCATGCTTCTCCTTCCCTTC 636

Db 698 ACCTGGGTCAACCAAGGTGGTCATCTTTGAAGCCTGCATGTTCATCTCCTGGTGGGACCACTC 757

QY 637 TCCATCATCATGGCCCTCTATGCTTGCATCTCTAGGGGCTGTGCTCCGAATACGCTCTGCT 696

Db 758 TGCCGTGTGCTGTCTCTCCTACTCACACATCTCTGGGGGCACTCTGAGGATCCAGTCTGGG 817

QY 697 CAGGCGTGGAAAAAGCCCTGGCCACCTGCTCTCTCCCACTTAACAGCTGTCAACCCCTTTC 756

Db 818 GAGGCGCGCAGAAAGGCTTCTCCACCTGCTCTCCCACTCTGCGTAGTGGGACTCTTC 877

QY 757 TATGGGCGAGCCATGTTCATGTACCTGAGGCCCTAGGCGCTAACCGGCCCTAGGCCATGAC 816

Db 878 TTTGGSAGCGCCATCGTCATGTATACATGGCCCCCTAAGTCCCGCCATCTGAGGAGCAGCAG 937

QY 817 AAGGTGCGCTCTATCTTCTACACAGATCCTTACTCCCATGCTG---AACCCCTCATTTA 872

Db 938 AAGGTCTTTTCTTATTTTACAGTTCTTCAACCCCGATGCTTAACCCCTCATTTA 997

QY 873 C-AGCTTGAGGAATGGGAGG 892

Db 998 CAACCTGAGGAATGTAGAGG 1018

RESULT 3  
 US-08-467-947A-1  
 : Sequence 1, Application US/08467947A  
 : Patent No. 6090575  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: LI, YI  
 : APPLICANT: CAO, LIANG  
 : APPLICANT: NI, JIAN  
 : APPLICANT: GENTZ, REINER  
 : APPLICANT: BULT, CAROL J.  
 : APPLICANT: SUTTON III, GRANGER G.  
 : APPLICANT: ROSEN, CRAIG A.  
 : TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
 : TITLE OF INVENTION: Coupled Receptor GPR1  
 : NUMBER OF SEQUENCES: 30  
 :  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 : STREET: 1100 NEW YORK AVE., NW, SUITE 600  
 : CITY: WASHINGTON  
 : STATE: DC  
 :  
 : COUNTRY: USA  
 : ZIP: 20005  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: FLOPPY DISK  
 : COMPUTER: IBM PC COMPATIBLE  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/467, 947A  
 : FILING DATE: 06-JUN-1995  
 : CLASSIFICATION: 435  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US95/04079  
 : FILING DATE: 30-MAR-1995  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: STEEFE, ERIC K.  
 : REGISTRATION NUMBER: 36,688

```

; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116..1003
;
US-08-467-947A-1

```

Query Match	25.8%;	Score 245;	DB 3;	Length 1713;
Best Local Similarity	56.8%;	Pred. No. 3.2e-61;		
Matches 489;	Conservative 1;	Mismatches 366;	Indels 5;	Gaps 2;
QY 37	TTCCTCCCTCTGGGCATCTTTCCACAGCCAGACTGACCTTGTCCTCTCTGACAGTT	96		
Db 158	TTCCCTCTACTGGGATTTCTCCCTGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTC	217		
QY 97	ATGGTGGCTCTCACAGTGGCCCTCTGTGGGAATGTCCTCATCTTCTCATCTACCTG	156		
Db 218	TCCCTGTCTATGTCTTACCCCTGCTGGGGAATGGAGCATCTCTGGGGCTCATCTCACTG	277		
QY 157	GACGCTGGACTTCACACCCCCCATGTACTTCTCTCAGCCAGCTCTCCCTCATGACCTC	216		
Db 278	GACTCCAGACTCCACACCCCCCATGTACTTCTCTCTCACACCTGGGCCGTGTCACATC	337		
QY 217	ATGTTGGTCTGTAAACATTTGTGCCAAGATGGCAGCAACTCTCTGTCTGGCAGAGAAGTCC	276		
Db 338	GCCTATGCTTCACACACAGTGGCCCCAGATGCTGTGAACCTCTCTGCATCCAGCCAAAGCCC	397		
QY 277	ATCTCCTTTGTGGGCTGTGGCATACAAATTTGGCTTTTGTCTCTCTGTGGGATCTGAG	336		
Db 398	ATCTCCTTTGTGGGCTGTGGCATACAAATTTGGCTTTTGTCTCTCTGTGGGATCTGAG	457		
QY 337	GGGCTCTTGTGGGACTCATGGCTTATGACCACTACGTGGCCGTATAGCCACCCACTTTCAC	396		
Db 458	TGCCCTCTGTTGGTGTGATGTCTTACGATCGGTACGTGGCCATCTGCCACCCCTCTCCGA	517		
QY 397	TATCCATCCTCATATGATCAGAGGGTCTGTCTCCAGATTTACTGGGAGCTCCCTGGGCTTT	456		
Db 518	TATTTTCATCATATGACCTTGAAAGTCTGCATCCTCTGGGCATCACTTCTCTGGACATGT	577		
QY 457	GGGATATAGATGAGTATTCAGATGGTGGCAGCCATGGGCTTACTTACTGTGGCTCA	516		
Db 578	GGCTCCCTCTGGCTATGTGTCATGTAGCCCTCATCTTAAGACTGCCCTTTTGTGGGCTT	637		
QY 517	AGAGCGTGGATCACTTTTCTGTGAGGTACAGACCTTATTTGAACCTGGCCGTGCAGAC	576		
Db 638	CGTGAATCAACACACTTCTTCTGTGAAATCCCTGTCTCTCAGGCTGGCCGTGTCTGAT	697		
QY 577	ACTTCCTTTTGTGACACCCCTCTTGTGCTTGTCTGTCTTCTCATCTTCTCTCCCTTTC	636		
Db 698	ACCTGGCTCAACACAGTGTGTATCTTTGAAGCCCTCATGTTTATCTCTGTGGACCACTC	757		
QY 637	TCCATCATCATGGCCCTCTATGCTTGACATCCCTAAGGGGCTGTGCTCCGATATAGCTCTGCT	696		
Db 758	TGCTGTGTCTGTCTCTACTACACATCTCTGGGGGATCTCTGAGGATCCAGCTCTGGG	817		
QY 697	CAGGCTTGAAAAAAGCCCTGGGCACTGTCTCTCCACCTTAACAGCTGTACACCTCTTC	756		
Db 818	GAGGCGCCAGAAAGCCCTTCTCCACCTGTCTCTCCACCTCTGTGCTAGTGGGACTCTTC	877		
QY 757	TATGGGCGACCATGTTATGTAAGGCTTGAAGGCTTACCGGGCCCTAGCCATGAC	816		
Db 878	TTTGSAGCGCCATCTCATGTACATGGCCCTTAAGTCCCGCCCATCTGAGAGACAGCAG	937		
QY 817	AAGGTGGCCCTTATCTTCTACACAGTCTTACTCCATGCTG---AACCCTCATTTTA	872		



TELEPHONE: 312-616-5600  
TELEFAX: 312-616-5700  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 966 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-748-506-8

Query Match 24.6%; Score 233.6; DB 3; Length 966;  
Best Local Similarity 54.7%; Pred. No. 4.9e-58;  
Matches 464; Conservative 0; Mismatches 384; Indels 0; Gaps 0;

```
OY 75 CCTGTCTCTCTCTCTGACGATGATGTTGGTCTTCACAGTGGCCCTGTGGGAATGTCCT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 CTTCCTCTCTCTACCCCTCATCTCTCTCATGTTCTTAGTATCACTAACAGAAATGCTCT 149
OY 135 CCTCATCTCTCTCATCTACTGAGCGCTGAGCTTCACACCCCATGTACTTCTTCTCAG 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 CATAGCCCTTGTCTATTGTGACAGTCCATCTCTACACACCCCATGTACTTCTTCTGCC 209
OY 195 CCAGCTCTCCCTCATGAGACCTCATGTTGGTCTGTAACTGTGCCAAGATGCCAGCCAA 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CAACTGTCTCTCTGAGATGCGTATACCTGCTCTGTCATACCCCAAGATGCTGCAGAG 269
OY 255 CTTCCTGTCTGCGAGGAGATCCATCTCTTGTGGCTGTGCATACAAATGGCTTTT 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 TCTGTGAGTGAAGCCCGAGAGATCTTTCAGGTGGAGTGGCCACAGATGTTTCTT 329
OY 315 TGTCTCTTGTGGGATGAGGGGCTCTTGTGGAGTCAATGCTTATGACCACTACGT 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 CATATCTTGTGATTAACAGAGTGTGCTGCTATTTGGCAGCATGGCCTTTCACCGTATAT 389
OY 375 GGGCGTTAGCCACCACCTTCACTATCCCATCTCATGAATCAGAGGGTCTCTCCAGAT 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 GGGTATATGTTCCCACTCACTATGCAACCCGAATGAGTGTGAGTATGTGCCACTT 449
OY 435 TACTGGAGCTCTGCGGCTTTGGGATTAATAGATGAGTTCAGATGTTGAGCCAT 494
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 GGCATTTGTTTCATGGGATGAGATGCAATGAGTCTGGGACAGACCAATTTTATTTT 509
OY 495 GGGCTTACCTTACTGTGCTCAAGAGCGTGAATCACTTTTCTGTGAGGTACAAGCTTT 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 CTCCTTGAACCTTCTGTGACCTGTGAGATGACCACTTCTGTGATCTTCCACCTCT 569
OY 555 ATGAAAGCTGGCTGTGACAGACACTTCCCTTTTGTGACACCTCTCTTGTCTGTGT 614
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 CCTGGACCTTGGCTGTGATACATCCCAATTTGAGCTGACCATCTTGTGTAGTTGT 629
OY 615 CTTCATGCTTCTCTCTCTCTCATGATGAGCTTCCATGCTTGTGATGCTTACAGGCG 674
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 CCTTGCATATCTAGCCCTTTTGTGATCATTTATTTATGTAGAAATCTCTGTTGC 689
OY 675 TGTGCTCGAATAGCGTCTGTGAGCGCTGGAAGGAGGCTGAGCTGCTCTCCCA 734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 AGTGTGTGATGCTTCACTGAGGGGCGCACAAAGCCCTTCAACCTGTCTCTCCA 749
OY 735 CCTAAGCTGTCAACCTCTTCTATGAGGAGCAGCATGTCTATGTAAGGCGCTAGGCG 794
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 750 CCTACTGTAGTACACACTCTTTATGCTCAGGATCTGTACCTATTTGAGGCGCTAGTC 809
OY 795 CTACCGGCGCCCTAGCCATGACAAAGTGGCTCTATCTTCTACACAGTCTTACTCCAT 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 810 TAGCCACTCACAGAAATGACAAACTTGGCCCTTCTTACACAGCAGTACATCCAT 869
OY 855 GCTGAACCCCTCATTTACAGCTTGAAGAAATGGGAGTGTGAGGCGACTGAGAAAGG 914
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 870 GTTGAACCTTATCATATAGTTTAAGGAACAAGATGTCAAGGACAGCACTGAGAAAGAT 929
OY 915 GCTGAGC 922
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 930 TCTGGCCC 937

# RESULT 6

US-08-748-506-5

; Sequence 5, Application us/08748506

; Patent No. 6159707

; GENERAL INFORMATION:

; APPLICANT: Ronnett et al.

; TITLE OF INVENTION: NOVEL SPERM RECEPTORS

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leydig, Volt & Mayer, Ltd.

; STREET: Two Prudential Plaza, Suite 4900

; CITY: Chicago

; STATE: IL

; COUNTRY: US

; ZIP: 60601-6780

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/748,506

; FILING DATE: 08-NOV-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 60/033,751

; FILING DATE: 09-NOV-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; REFERENCE/DOCKET NUMBER: 74940

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-616-5700

; TELEFAX: 312-616-5600

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 966 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-748-506-5

Query Match 24.6%; Score 233.2; DB 3; Length 966;

Best Local Similarity 54.6%; Pred. No. 6.4e-58;

Matches 463; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

```
OY 75 CCTGTCTCTCTCTCTGACGATGATGTTGGTCTTCACAGTGGCCCTGTGGGAATGTCCT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 CTTCCTCTCTCTACCCCTCATCTCTCTCATGTTCTTAGTATCACTAACAGAAATGCTCT 149
OY 135 CCTCATCTCTCTCATCTACTGAGCGCTGAGCTTCACACCCCATGTACTTCTTCTCAG 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 CATAGCTCTTGTGATTTGATACAGTCCATCTCTTACACACCCCATGTACTTCTTCTGCC 209
OY 195 CCAGCTCTCCCTCATGAGACCTCATGTTGGTCTGTAACTGTGCCAAGATGCCAGCCAA 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CAACTGTCTCTCTGAGATGCGTATACCTGCTCTGTCATACCCCAAGATGCTGCAGAG 269
OY 255 CTTCCTGTCTGCGAGGAGATCCATCTCTTGTGGCTGTGCATACAAATGGCTTTT 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 CTTGTGAGTGAAGCCAGAGAGATCTCTGGGAGGAGTGTGCCACAGATGTTTITTTT 329
OY 315 TGTCTCTTGTGGGATGAGGGGCTTGTGGGAGTCAATGCTTATGACCACTACAGT 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 CGCATTTTGTGATTAACAGTGTGCTGCTTATTTGGCAGCATGGCCTTTGACCGCTGAT 389
OY 375 GGGCGTTAGCCACCACCTTCACTATCCCATCTCTCATGAATCAGAGGGTCTGTCTCAGAT 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 GGCATATGCTCCCACTCCACTATGCAACCCGAATGAGTGTGAGGATGTGCTTATTT 449
```



OY	435	TACTGGAGGCTCCTGGGCCCTTTGGATATATAGATGGAGTGTATTCAGATGGTGCCACCAT	494
Db	450	GCGAATTGTTTCATGGGGGAATGGGATGCATAGTAAGTCTGGGACAACCAATTTATTTT	509
OY	495	GGGCTTACCCTTACTGTGGCTCAAGGAGCGTGGATCACTTTTTCTGTGAGGTACAAGCTTT	554
Db	510	CCTCTTGAACCTCTGTGAGCCCTGTGAATAATAGACCACCTTCTGTGACCTTCCACCTCT	569
OY	555	ATTGAAGCTGGCGCTGTGACAGACACTTCCCTTTTGAACCCCTCCCTTTGCTGCTGTGT	614
Db	570	CCTGGCACCTTGCCTGTGAGATATACATCCCAAACGAGCGTGCACCTTTGTGGTAAGCAGT	629
OY	615	CTTCACTGCTTCTCCTTCCCTTCTCCATCATCATGCGCTCCTATGCTTGACATCCTAGGGC	674
Db	630	CCNCTGCATATCTAGCCCAATTTTGTCTATCATTTATTTCTATGCAAAATTTCTATTTGC	689
OY	675	TGTGCTCCGAATPACGCTCTGCTCAGCGCTGGAAAAAGCCCTGGCCACCTGCTCTCCCA	734
Db	690	AGTGTNTGTATGCTTCACTTGAGGGCGCCATAAAGCTCTTTCACCTGTTCGTCTCA	749
OY	735	CCTAACAGCTGTACACCCCTTCTATGSGGACGCCATGTCATGTACTTGAGGCGCTAGGCG	794
Db	750	CCTAATTGTAGTCACACTTTTATGCTCAGCATGTATTACCTATTGTGAGGCCCAAGTC	809
OY	795	CTACCGGCGCCCTAGCCATGACAAGAAGTGCGCTCTATCTTCTACAGAGTCCCTTACTCCAT	854
Db	810	TAGCCACTCACCGAAGATGACCAAAATTCCTTGCGCCCTCTTCTACACAGTAGTAGACATCCAT	869
OY	855	GCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGGCCACTGAGAAGGG	914
Db	870	GCTGAACCCCTATCATCTATATAGTTTAAGGAACAAGGAGTCAAGGACGACCTGGAAGAAC	929
OY	915	GCTGGACC 922	
Db	930	TCTGGGCC 937	

RESULT 7  
US-08-748-506-6  
Sequence 6, Application US/08748506  
Patent No. 6159707  
GENERAL INFORMATION:  
APPLICANT: Ronnett et al.  
TITLE OF INVENTION: NOVEL SPERM RECEPTORS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Volt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: US  
ZIP: 60601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748, 506  
FILING DATE: 08-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/033, 751  
FILING DATE: 09-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
REFERENCE/DOCKET NUMBER: 74940  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5600  
TELEFAX: 312-616-5700  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 966 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

```

Query Match	24.08;	Score 227.4;	DB 3;	Length 966;
Best Local Similarity	54.18;	Pred. No. 3e-56;		
Matches 459;	Conservative 0;	Mismatches 389;	Indels 0;	Gaps 0;

QY	75	CCCTTGTCCTCTTCTCTGCAAGTTATGGTGTCTTCACAGTGGCCCCCTCTGTGGGAATGCTC	134
Db	90	CTTCCCTCGTTTCACCCCTCATCCTTCTCTATGTTCTTAGTATACACTAACAGGAATGCTCT	149
QY	135	CCTCATCTTCCATCATCTFACCCTGGACCGCTGGACTTCACACCCCCCATGTACTTCTTCTCAG	194
Db	150	CATAGCCCCTTGCTGNTTGTACCAAGTCCATCTCTACACACCCCCCATGTACTTCTTCTGGC	209
QY	195	CCAGCTCTCCCTCATAGGACCTCATGTGGTGTCTGTAAATTTGTGCCAAGAATGGCACCCA	254
Db	210	CAACTTGTCTCTCCTGGAGATTGGCTATACTTGTCTGTATACCAAGAATGCTGCAGAG	269
QY	255	CTTCCCTGTCGTGGCAGGAAGTCCATCTCCTTTGTGGCGTGTGGCATACAATTGGCCTTTT	314
Db	270	CCTTGTGAGTGAGGCCAGAGAGATCTCTAAGGAGGGATGTGCCACACAGATGTTTTTCTT	329
QY	315	TGTCCTCTCTTGTGGGATCTGAGGGGCTCTTGTCTGGGACTCATGGCTTATGACCCTACGT	374
Db	330	CACATTTTGTGGCATAACTGAGTGTGCTTATTTGGCAGCCATGGCCTTTGACCGCTGCAT	389
QY	375	GCGCGTTAGCCACCCACCTTCATCTATCCATCCTCATGAAATCAGAGGGTCTGTCTCCAGAT	434
Db	390	GCGCATATGCTCCCCACCTCCACTATGCAACCCGAATGAGTGTGAGGATGTGCCCATTT	449
QY	435	TACTGGAGCTCCTGGGCTTTGGGAAATAGATGAGTGAATCAGATGGTGCAGCCAT	494
Db	450	GCGCAATTGTTTCATGGGGATGGGATGAGTATAGTCTGTGGACAGACCAAATTNNATTN	509
QY	495	GCGCTTACCTTACTGTGTGCTCAAGAGCGGTGATCACTTTTCTGTGAGGTGTCAAGCTTT	554
Db	510	CTCCTTGAACCTCTGTGACCTTGTGAGATAGACCACTTCTTCTGTGACCTTCCACCTCT	569
QY	555	AATGAAGCTGGCCTGTGACAGACACTTCCCTTTTGTGACACCCCTCCTTGTGCTGTGT	614
Db	570	CTGTGCACCTGCTGTGTGATACATGCCAAAACGAGGCTGCCACTTGTGTGCGACGAAT	629
QY	615	CTTCATGCTTCTCCTTCTCTCCATCATCATGCGCTCTATGCTTGACCTAGGGGC	674
Db	630	CCCTGTATATCTAGTCCATTTTGTGATCCTTTATTTCTTATGTCAAGAACTCTCGTGC	689
QY	675	TGTGCTCCGAATACGCTGTGCTCAGGCTTGAAAAAAGCCCTGGCCACCTGCTCTCCCA	734
Db	690	AGTGCTGTGATGCTTCACTGAGGGGCGCCATAAAGCTCTTCCACCTGTTCCTCACA	749
QY	735	CCTAACAGCTGTACCCCTTCTTATGGGGCAGCAATGTTCAATGTACCTGAGGCTTAGGG	794
Db	750	CCACTTGTATACACACTCTTTTATGGCTGTGTCTTACCTAATTTGAGGCCCAAGTC	809
QY	795	CTACCGGGCCCCTAGCCATGACAAGTGGGCTCTATCTTCTACACAGTCTTACTCCCAT	854
Db	810	TAGCCACTCACAGGAATGACAACACTTGGCCCTCTTCTACACAGCAGTGCATCCAT	869
QY	855	GCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGGCCTGAGGAAGG	914
Db	870	GCTGAACCCCTATCATCTTACAGTCTTAAGAACAAAGGAAGTCAAGCCAGCACTGAGAGAAC	929
QY	915	GCTGAGCC 922	
Db	930	TCTGAGCC 937	

RESULT 8  
US-09-668-680-9

```
Sequence 9, Application US/09668680
Patent No. 6436703
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Wang, Jian-Rui
APPLICANT: Xue, Aidong J.
APPLICANT: Xu, Chongjun
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6436703el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 790CIP2A
CURRENT APPLICATION NUMBER: US/09/668,680
CURRENT FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 13
SOFTWARE: pt_genes Version 2.0
SEQ ID NO 9
LENGTH: 1080
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (178)..(1080)
US-09-668-680-9
```

```
Query Match 21.9%; Score 208; DB 4; Length 1080;
Best Local Similarity 52.3%; Pred. No. 1.3e-50;
Matches 460; Conservative 0; Mismatches 420; Indels 0; Gaps 0;

QY 37 TTCTTCCTCTTGGGACATCTTTCCACAGCCAGACTGACCTTGTCTCTCTCTGACGTT 96
Db 112 TTCTTCCTCTTGGGACATCTTTCCAGAGATCCAGAACTGACGCCCCCTCTCCCTGGCTGTC 171
QY 97 ATGCTGCTCTTACAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCT 156
Db 172 CTGTCATGTACCTGCTGACGCGGTGAGGAGACCTGCTCATCTCTCTCTCTCTCTCTCTCT 231
QY 157 GACGCTGACCTTCAACACCCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 216
Db 232 GACTCCACCTCCACACCCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291
QY 217 ATGCTGCTCTTACAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Db 292 GGTTCACCTCGGCGCATGTTCCCAAGATGATGTGACATGACATGCGCATAGCAGAGTC 351
QY 277 ATCTCTTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 336
Db 352 ATCTCTTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 411
QY 337 GGGCTCTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 396
Db 412 GACATGCTCTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 471
QY 397 TATCCATCTCTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCT 456
Db 472 TATCCATCTCTGTGGGCTGTGGCATCAAAATGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 531
QY 457 GGGATATAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 516
Db 532 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
QY 517 AGGAGCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
Db 592 GTGGAATCTCCAAATTTTGTCTGTGACCCCATCTCAACTTCTCAACTTGGCTGTCTGTGAC 651
QY 577 ACTTCCTTTTGTGACACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 636
```

```
Db 652 AGTGCATCATAGCATATATATATATATATATATATATATATATATATATATATATATATAT 711
QY 637 TCCATCATCATAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
Db 712 TCAGGATCTCTTTGTCTTACGTAACAATGTCCCTCCATCTTAAGAATTCATCATCA 771
QY 697 CAGGCTTGAAAAAGCCCTGCGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 756
Db 772 GATAGGAAGTCTAAAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 831
QY 757 TATGGGCGAGCATGTTTATATATATATATATATATATATATATATATATATATATATATATAT 816
Db 832 TATGGAACAGCATGTTGCGGTGTAACCTGACTTCACTGTGTACACACCCCGAAGATGT 891
QY 817 AAGTGGCCTCTATCTTCTTACAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 876
Db 892 GTGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
QY 877 TTGAGGAATGGGAGGTGATGGGGGCACTGAGGAAGGGC 916
Db 952 CTGAGAAATAGGACATTTCAAAAGTGCCCTGTGAGAGGCTGC 991
```

```
RESULT 9
US-09-085-371-5
Sequence 5, Application US/09085371
Patent No. 6218358
GENERAL INFORMATION:
APPLICANT: Firestein, Stuart
APPLICANT: Zhao, Haiqing
TITLE OF INVENTION: Functional Expression of, and Assay for, Functional Cellular
FILE REFERENCE: P0151US2 / 09805059
CURRENT APPLICATION NUMBER: US/09/085,371
PRIOR APPLICATION NUMBER: US 08/891,243
PRIOR FILING DATE: 1998-05-19
PRIOR APPLICATION NUMBER: US 60/045,961
PRIOR FILING DATE: 1997-07-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 900
TYPE: DNA
ORGANISM: Rat
US-09-085-371-5
```

```
Query Match 21.7%; Score 205.4; DB 4; Length 900;
Best Local Similarity 54.3%; Pred. No. 6.7e-50;
Matches 444; Conservative 0; Mismatches 361; Indels 12; Gaps 1;

QY 80 TCTCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 139
Db 80 TACTATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 139
QY 140 TCTCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 199
Db 140 TTATAGCAATTAAGAACACCAACCTCCACAAACCAACCAATTTTCTTGGCTAATA 199
QY 200 TCTCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 259
Db 200 TGTCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 259
QY 260 TGTCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 307
Db 260 TTGCTCTCTCTGACGTTATGTTGTTCTTCAAGTGGCCCTCTGTGGGAATGCTCTCTCTCTCA 319
QY 308 GCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 367
Db 320 ACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 379
QY 368 ACTAGTGGCGGTAGGACACCCACTTCACTATCCATCCATCATGATCATGAGGCTCTGTC 427
```

```
Db 380 GCTATGTGGCTATCTGTCTCCACTCCACTACCCCGTCAATGTGCTAGCCGGCTATGTG 439
QY 428 TCCAGATTACTGGAGCTCCCTGGCCCTTGGGATAATAGATGGAGTGAATTCAGATGGTGG 487
Db 440 TGCAGATGGCAGCTGATCTCTGGGCTGGAGGTTTGGTATCTCCATGGTTAAAGTTTCC 499
QY 488 CAGCCATGGGCTTACCTTACTGTGGCTCAGAGAGCGGTGATCACTTTTCTGTGAGGTAC 547
Db 500 TTATTTCTGGCCTGTCTTACTGTGGCCCAACACCATCAACCACTTTTCTGTGATGTGT 559
QY 548 AAGCTTATGAGCTGGGCTGTGACAGACTTCCCTTTTGAACCCCTCTCTTGTCTT 607
Db 560 CTCATTGTCTCAACCTGTGATGCACTGACATGTCCACAGCAGAGCTTACAGACTTGTCC 619
QY 608 GCTGTCTTCTCATGCTTCTCCTTCCCTTCTCCATCATCATGAGCTTCTTGTGATCC 667
Db 620 TGGCCATTTTATTTCTGTGGGAGCCGCTCTGTCTACCTGGGGCATCTTACATGGCCATCA 679
QY 668 TAGGGGCTGTCTCCGATACGCTCTGCTCAGGCTGGAAGAAAGCCCTGGCCACCTGCT 727
Db 680 CAGGTGCTGTGATGCGCATCCCTCAGCTGCTGGCCGCAATAAGCCTTTTCAACCTGTG 739
QY 728 CTCCCACTTAACAGCTGTCACCCCTCTTCTATGGGGCAGCCATGTCTATGATCCTGAGGC 787
Db 740 CCTCCCACTCACTGTGTGATCATCTTCTATGACAGCCAGTATTTCTATCTATGCCAGGC 799
QY 788 CTAGCGCTACCGGCCCCCTAGCCATGACAGAGGTGGCCCTCTATCTTCTACACAGTCCCTTA 847
Db 800 CTAGGCACTCTCAGCTTTTGACACCAACAAGCTGGTCTCTGTACTCTACGCTGTCTATG 859
QY 848 CTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAA 884
Db 860 TACCGTGTTCATCCCATCATCTACTGCTTGGCGAA 896
```

## RESULT 10

```
US-09-668-680-11
; Sequence 11, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6436703el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2A
; CURRENT APPLICATION NUMBER: US/09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 11
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1091)
US-09-668-680-11
```

```
Query Match 21.5%; Score 203.4; DB 4; Length 1297;
Best Local Similarity 53.3%; Pred. No. 3e-49;
Matches 429; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
```

```
QY 124 GGAATGTCCCTCCTCATCTTCTCCTCATCTGAGCGCTGGACTTACACCCCCCATGTAC 183
Db 246 GGAACCTGCTCATCTCTGGCCATCAGCCCTGACTCCACCTCCACACCCCATGTAC 305
QY 184 TTCTTCTCAGCCAGCTCTCCCTCATGAGCCTCATGTGTGTGTGTATGATGTGCAAG 243
Db 306 TTCTTCTCTCCACCTGCTCCCTTGGCTGACATCGGTTCACCTCCACACCGTCCCAAG 365
QY 244 ATGGCAGCAACTCTCTCTCTGAGAGAGTCCATCTTCTTGTGGCTGTGGCATACA 303
Db 366 ATGATTGTGACATGCAACTCAGAGCAGATCATCTCTATGAAGGCTGCTGACTCAG 425
QY 304 ATGGCTTTTGTCTCTCTTGTGGATCTGAGGGCTCTTGTGCTGAGCTCATGGCTTAT 363
Db 426 ATGCTTTTGTCTCTCTTGTGATGTGATGATGATGATGATGATGATGATGATGATGAT 485
QY 364 GACCACCTAGCTGGCCGTTAGCCACCCACTTCACTATCCATCTCTCATGAATCAGAGGTC 423
Db 486 GACCGGTTTGTGGCCATCTGTACACCCCTGACCTACCGAATCATCATGAACCCACGCTC 545
QY 424 TGTCTCCAGATTACTGGAGCTCTCTGGGCTTTGGGATTAATAGATGAGTGAATCAGATG 483
Db 546 TGTGGCTTCTTAATCTGTGTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 605
QY 484 GTGGCAGCATGGGCTTACCTTACTGTGCTCAAGAGCGTGGATCATTCTTCTGTAG 543
Db 606 TTGATTATGTATACAGCTACCTGCTTCAAGAGATGTGACATTTCTAATTTCTTCTGTAC 665
QY 544 GTACAAGCTTATTAAGCTGGCTGTGACAGACACTTCTTGTGACACCTCTCTT 603
Db 666 CTTCTCACTCTCTCCACCTTAGGTGTTCGACACCTTCATCATGAATGTCATATAT 725
QY 604 GCTTGTGTCTTCTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
Db 726 TTCAATGGTGCATATTTGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 785
QY 664 ATCTAGGGGCTGTGCTCCGATATAGCTCTGCTCAGGCTTGAAAAAGCCCTGGCCACC 723
Db 786 ATTGTTCCCTCATTTCTGAGAGTTCCACATCAGATGGGAAGTATAAGCCCTTCTCCACC 845
QY 724 TGCTCTCCCACTTAACAGCTGTACACCTCTTCTATGAGGGCAGCCATGTCATGTC 783
Db 846 TGTGGCTCTACCTGGAGTGTGTTGCTTATTTATGGAACAGGGCTTGTAGGTCCTC 905
QY 784 AGGCTAGGGCTACCGGGCCCTAGCCATGACAAGTGGCTCTATCTTCTACACAGTC 843
Db 906 AGTTAGCTGTGTACCATCTCCCAAGAGAGTATGAGTGGCTTACATGATGTACACTGTG 965
QY 844 CTACTCCCATGTGTAACCCCTCATTTACAGCTTGAAGGAATGGGAGGTGATGGGCA 903
Db 966 GTCACCCCATGTGTAACCCCTCATTTACAGCTTGAAGGAACAGGACATTCAAAGTCC 1025
QY 904 CTGAGGAAGGGCTGACCCGCTGCA 928
Db 1026 CTGTGAGGCTGCATGCGAGAATCA 1050
```

## RESULT 11

```
US-09-668-680-10
; Sequence 10, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6436703el Nucleic Acids and
```



```

QY 337 GGGCTCTTGCTGGGACTCATGGCTTATATACCACTACGTRGGCCGTAGACCACCACTTAC 396
Db 352 TTTTTCCTTTTGGCTGCGTTATCCCTGGACCGCTTTCGTGGCCATCTGCAAACTCTACAT 411
QY 397 TATCCCATCCTCATGTAATCAGAGGGTCTGTCTCCAGATTAATCTGGAGCTCCCTGGCCCTTT 456
Db 412 TATCCCAACCATCATGAGCCCAAGGATGTGCTCTCTTCGTTACTGTCTGTTATTTTGG 471
QY 457 GGGATTAATAGATGAGATGATTCAGATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCA 516
Db 472 GGCTTCCTCTTCATGAGCCAGTCCAGTGTGATGCTTTTCCAAAGACATTTTACTGTGTCCA 531
QY 517 AGGAGCGTGATCATCTTTTCTGTGAGGTACAAAGCTTATTTGAAGCTGGCCCTGTGCAGAC 576
Db 532 AACATATATCCTCCTACCTTTTCTGTGATTTTGGACCACCTGGCAAACTCTCTCCTGTTCAGAA 591
QY 577 ACTTCCCTTTTGGACACCCCTCCTTTGGCTGTGCTGTGCTTCATGCTTCTCCTTCCCTTC 636
Db 592 ACCAGGCTCTATTGAGATGCTGTTTTTTTACCCCTTGCTGTAAATGTGCTTTTGGCTTCCCTTT 651
QY 637 TCCATATCATGAGCGCTTCCTATGCTTGCACTCCTAGGGGCTGTGCTCCGAATACGCTCTGCT 696
Db 652 CTTATAGCCATCTTTGCATACAGCAATATAGTAGTACACCATAGTGAGAGACTCCCTTCAGCC 711
QY 697 CAGGCGCTGAAAAAAGCCCTGGCCACCTGCTCCTCCACCTAACAGCTTCACCCCTTTC 756
Db 712 AGGAGCGACAGAGAGAGCTTTTTCACCTGCTCCTCTCTCATCTCATTTGCTCTCTCTAATG 771
QY 757 TATGGGCGACCATGTTTCATGTACCTGAGGCCCTAGGGGCTTACCGGGCCCTAGCCATGAC 816
Db 772 TATGGCAGCTGTGCATTTATATACCTGAAGCCAAAGCAGAGAAGCAGAGTGACACCAAC 831
QY 817 AAGGTGGCCTCTATCTTCTACACAGTCCCTTACTCCCATGCTGAACCCCTCATTTACAGC 876
Db 832 AGAGAGGCTGCTCTTGTGAACATGTTGTGACACCCCTTCTGAACCCCTGTCACTACACC 891
QY 877 TTGAGGAATGGGAGGTGATGGGGGCACTGAGGAAGGGGCTGGACCG 923
Db 892 CTGCGCAACAAGCAGGTCCACCAGGCTCTCAGGAGTGTCTGTCCAG 938

RESULT 13
US-09-605-785-526
; Sequence 526, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605,785
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA

```

Query Match	12.9%	Score 122.4	DB 4	Length 963
Best Local Similarity	49.0%	Pred. No. 6,7e-26		
Matches 327; Conservative	0	Mismatches 341	Indels	Gaps
				0;
111 AGTGCCCTCTGTGGGAATGCTCCATCTTCTCATCTACCTGAGCGCTGACCTTCA	170			
105 AGTGCAATGTTGGAACTGCATCGTGCTTCATCTAAGAGCAACGACGCTGCA	164			
171 CACCCCATGTACTTCTTCTCAGCAGCCTCTCCATATGACCTCATGTTGTCTGTA	230			
165 CGCTCCGATGTACTTCTTCTGCAATGCTTGACGACCATTGACCTGACCTATC	224			
231 CATTGCCCCAAGATGGCAGCCAACTTCTGCTGCGAGAGATCCATCTCTTGTGG	290			
225 CACCATGCTTAAGATCTTCCCTTCTGCTTGTGATTCCGAGAGATTAGCTTGA	284			
291 CTGTGCATACAAATGCGCTTTTCTCTCTCTGTTGGGATCTGAGGGGCTTGTG	350			
285 CTGTCTTACCCAGATGTCTTATATGATGCCCCCTGACCATTTGAATCCACAT	344			
351 ACTATGCGCTTATGACCACTAGCTGCGGCTTAGCCACCACCTTCACTATCCAT	410			
345 GGCATGCGCTTGTGACCGTTATGTGGCATCTGCCACCACCTGCGCATGTGAG	404			
411 GAATCAGAGGCTCTGTCTCCAGATTAATGAGAGCTTGGGCTTGGATTAATAG	470			
405 CAACATATACAGTAACAGCCAGATGCGATCGTGCTGTGTGCGCGGATCCCT	464			
471 AGTATTCAGATGTGTGCGAGCCATGGGCTTACCTTACTGTGTGCTCAAGAG	530			
465 TTTCACACTGCTCTGTGATCAAGCGGCTGGCTTCTGACCTCAATGTCTTCG	524			
531 CTTTCTGTGAGGTACAAAGCTTTATGAAGCTGGCCTGTGCAGACACTTCC	590			
525 CTCTTATTTGTTCACCAAGATGTAATGAAGTTGGCTATGCAGACACTTGG	584			
591 CACCTTCTTGTGCTGTGTCTCATGCTTCTCTCTCTCTCTCTCTCTCTCT	650			
585 GGTATATGCTCTTACTGCAATTCGTGTGTCATGGCGGTGAGCTAATGTTCA	644			
651 CTCTTATGCTTGCATCCTAGGGGCTGTGCTCCGAATACGCTCTGTCTGAA	710			
645 GTCTTATTTTCTGATATATAGCAAGGTTTGCACAGCTTCCAACTCAGAG	704			
711 AGCCCTGCGCACCTGCTCTCTCCACCTTAACAGCTGTACACCTCTTCTAT	770			
705 GGCCTTGAACCTGTGTGTCAACATTTGTTGTGTACTCGCTTCTATGTGCA	764			
771 GTTCATGT 778				
765 TGGCTCT 772				

```

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSEQ for windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-526

```

Query Match	12.9%;	Score 122.4;	DB 4;	Length 963;
Best Local Similarity	49.0%;	Pred. No. 6.7e-26;		
Matches 327; Conservative	0;	Mismatches 341;	Indels 0;	Gaps 0;

QY	111	AGGGGCCCTGTGGAATGTCCCTCCTCATCTTCTTCATCTACCTGGAGCGTGACTTCA	170
Db	105	AGTGGCAATGTTTGGAACTGCATCGTGCTTCTCATCGTAAGACGGAACCGACGCTGCA	164
QY	171	CACCCCCATGTACTTCTTCTCAGCCAGCTCTCCCTCATGGACCCTCATGTTGGTCTGTAA	230
Db	165	CGCTCCGATGTAACTCTTCTCTCGCATGCTTGCAGGCATTGACCTGGCCTTATCCACATC	224
QY	231	CATGTGCCAAGATGCGCACCCAACTTCCTGTCTGGCAGGAAGTCATCTCCCTTGTGGC	290
Db	225	CACCATGCGCTAAGATCCTTGCCCTTTCCTGTGTTGATTCGCCGAGAGATTAGCTTTGAGGC	284
QY	291	CTGTGGCATACAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTTGTCTGGC	350
Db	285	CTGTCTTACCCAGATGTCTTATTATCATGCCCCCTCAGCCATTGAATCCACCATTCTGCT	344
QY	351	ACTCATGGCTTATGACCACATACTGCGCCGTTAGCACCCCACCTTACTATCCATCCTCAT	410
Db	345	GCCCATGGCCTTTGACCGTATATGTGGCATCTGCCAACCACCTGGCCATGCTGCAGTGT	404
QY	411	GAATCAGAGGGTCTGTCTCCAGATTACTGGAGCTCCTGGGCTTTGGGATATAGATGG	470
Db	405	CAACATACAGTAACAGGCCAGATTGGCATGTGGCTGTGGTCCGGGATCCCTCTTTTT	464
QY	471	AGTATTCAGATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGAGCGTGATCA	530
Db	465	TTTCCACATGCTCCTCGCTGATCAAGGGCTGGCCTTCTGCCACTCCATGTCTCTCGCA	524
QY	531	CTTTTCTGTGAGGTACAAAGCTTATTAAGAAGCTGGGCTGTGCAGACACTTCCCTTTTGA	590
Db	525	CTCCTATTGTGTCACCAAGGATGTAATGAAGTTGGCTTATGCAACACTTTGCCCAATGT	584
QY	591	CACCCCTCCTTTGCTGTGCTGTCTCATGTCTTCCCTCTCCATCATCATGGC	650
Db	585	GGTATATGTCCTTACTGCGCATTTCTGCTGTGCATGGGCGTGGAGCTAATGTTCATCTCCT	644
QY	651	CTCCTATGCTTGCACTCAGGGGCTGTGCTCCGAATACGCTCTGCTCAGGCCGTGAAAAA	710
Db	645	GTCCTATTCTGTATTAATACGAACGGTCTTGCAACTGCCCTTCCAAGTCAGAGCGGGCAA	704
QY	711	AGCCCTGGCCACCTGTGCTCCACCACTAACAGCTGCTCACCCCTCTTATGGGCGACCAT	770
Db	705	GGCCTTTGGAACCTGTGTGCACACACATGTGTGTGTACTCGCCTTCTATGTGCCACTTAT	764
QY	771	GTTCATGT 778	
Db	765	TGGCCTCT 772	

```

RESULT 15
US-08-465-980-1
; Sequence 1, Application US/08465980
; Patent No. 5756309
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
;

```

```

1  APPLICANT: Ruben, Steven M.
2  TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
3  NUMBER OF SEQUENCES: 8
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
6  ADDRESSEE: STUART & OLSTEIN
7  STREET: 6 Becker Farm Road
8  CITY: Roseland
9  STATE: New Jersey
10 COUNTRY: USA
11 ZIP: 07068
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PatentIn Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/465,980
20 FILING DATE: 06-JUN-1995
21 CLASSIFICATION: 536
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Ferraro, Gregory D.
24 REGISTRATION NUMBER: 36,134
25 REFERENCE/DOCKET NUMBER: 325800-446
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 201-994-1700
28 TELEFAX: 201-994-1744
29 INFORMATION FOR SEQ ID NO: 1:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 1474 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 MOLECULE TYPE: cDNA
36 FEATURE:
37 NAME/KEY: CDS
38 LOCATION: 274..1233
39
40 US-08-465-980-1

```

Query Match	12.9%;	Score 122.4;	DB 1;	Length 1474;
Best Local Similarity	49.0%;	Pred. No. 8.1e-26;		
Matches 327; Conservative	0;	Mismatches 341;	Indels 0;	Gaps 0;

QY	111	AGTGGCCCTCTGTGGGAATGTCCCTCCATCTTCCTCATCTACCTGGAGCGCTGACCTTCA	170
Db	378	AGTGGCAATGTGTGGAACCTGCATCGTGGCTTTCATCATGTAAGGACGGAAACGAGCCCTGCA	437
QY	171	CACCCCATGTACTCTCTCTCAGCAGCACTCTCCCTCATATGACCCTCATGTGTGCTGTAA	230
Db	438	CGCTCCGATGTACTCTTTCTCTGCATGCTTGCAGCCATGACCTGGCCCTATCCACATC	497
QY	231	CATTGTGCCAAGATGGCAGCCAACTTCTCTGTGCGCAGGAAGTCCATCTCCTTGTGGG	290
Db	498	CACCATGCCCTAAGATCCTTGCCCTTTCTGTGTTGATTCCGAGAGATTAGCATTTGAGGC	557
QY	291	CTGTGGCATACAAATGGCTTTTCTCTCTCTCTGTGGGATCTGAGGGGCTCTTGTGGG	350
Db	558	CTGTCTTACCAGATGTCTTTATTCATGCCCCCTCTAGCCATTGAAATCCACCATCTGCT	617
QY	351	ACTCATGGCTTATGACCACCTAGCTGGCGCTTAGCCACCACCTTCACTATCCCATCTCAT	410
Db	618	GGCCATGGCCCTTGGACCGTTATGTGGCATCTGCCACCACCTGCGCCATGCTGCAGTCT	677
QY	411	GAATCAGAGGCTCTCTCTCCAGATTACTGGAGCTCCTGGGCCCTTGGGATATAGATGG	470
Db	678	CAACATATACAGTAAACGCCACAGATTGGCATCGTGGCTGTGGTCCGGCGATCCCCCTTTT	737
QY	471	AGTGATTGAGATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGATCA	530
Db	738	TTTCCCACTGCCCTCTGCTGATCAAGCGGCTGGCCCTTCTGCCCACTCCAAATGTCTCTGCA	797
QY	531	CTTTTCTGTGAGGTACAAGCTTTATTTGAAGCTGGCCTGTGCAGACACTTCCCTTTTGA	590



Db	798	CTCCATATTGTGTCCACCACGAGATGTAATGAAGTGGCCCTATGCAGACACTTGGCCAATGT	857
QY	591	CACCCCTCCTTTTGGCTGTGTGTTTCATGCTTCTCCCTTCTCCATCATCATGGC	650
Db	858	GGTATATGGTCTTACTGCCATTCTGCTGGTCAATGGGGGTGGACGTAATGTTATCTCCTT	917
QY	651	CTCCTATGCTTGCATCCTAGGGGCTGTGCTCCGATAACGCTCTGCTCAGGCCTGAAAAA	710
Db	918	GTCCATATTTTCTGATATAATACGACGAGTCTTGCAACTGCCCTTCCAAGTCACAGCGGGCCAA	977
QY	711	AGCCCTGGCCACCTGCTCCTCCCACTAACAGCTGTCAACCCTTCTATGGGCGACCAT	770
Db	978	GGCCTTTGGAACCTGTGTGTACACACATTTGGTGTGTACTCGCCCTTCTATGTGGCACTTAT	1037
QY	771	GTTCATGT	778
Db	1038	TGGCCTCT	1045

Search completed: May 9, 2003, 04:49:05  
Job time : 80 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2003, 04:31:20 ; Search time 1765 Seconds  
(without alignments)  
667.470 Million cell updates/sec

Title: US-09-634-109D-1  
Perfect score: 948  
Sequence: 1 atgggaagatggtgaacca.....ggattggcagcagcactga 948

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	935.4	98.7	1172	9	US-10-023-601-7	Sequence 7, Appli
2	810.4	85.5	948	10	US-09-886-055-176	Sequence 176, App
3	526.6	55.5	669	9	US-10-259-430-19	Sequence 19, Appl
4	526.6	55.5	669	9	US-10-259-430-27	Sequence 27, Appl
5	525	55.4	669	9	US-10-259-430-23	Sequence 23, Appl
6	390.2	41.2	488	10	US-09-747-155-212	Sequence 212, App
7	388.8	41.0	1290	10	US-09-826-508-23	Sequence 23, Appl
8	361.8	38.2	975	10	US-09-886-055-430	Sequence 430, App
9	360.6	38.0	951	10	US-09-886-055-478	Sequence 478, App
10	345	36.4	948	10	US-09-886-055-436	Sequence 436, App
11	343.8	36.3	1047	10	US-09-886-055-434	Sequence 434, App
12	342	36.1	1110	10	US-09-886-055-458	Sequence 458, App
13	339.8	35.8	954	10	US-09-886-055-480	Sequence 480, App
14	330.2	34.8	1017	10	US-09-886-055-428	Sequence 428, App
15	326.4	34.4	972	10	US-09-886-055-164	Sequence 164, App
16	325.8	34.4	945	10	US-09-886-055-144	Sequence 144, App
17	325.8	34.4	996	9	US-09-898-586-17	Sequence 17, Appl
18	325.8	34.4	996	10	US-09-761-288-17	Sequence 17, Appl
19	324.8	34.3	1047	10	US-09-886-055-166	Sequence 166, App

20	324.2	34.2	994	9	US-09-898-586-15	Sequence 15, Appl
21	324.2	34.2	994	10	US-09-761-288-15	Sequence 15, Appl
22	320.4	33.8	996	10	US-09-886-055-476	Sequence 476, App
23	313.4	33.1	939	10	US-09-886-055-474	Sequence 474, App
24	308.4	32.5	1008	9	US-09-974-591-13	Sequence 13, Appl
25	307.8	32.5	939	10	US-09-886-055-422	Sequence 422, App
26	306.8	32.4	951	10	US-09-886-055-198	Sequence 198, App
27	306.8	32.4	1008	9	US-09-974-591-11	Sequence 11, Appl
28	303.2	32.0	1954	10	US-09-864-761-5098	Sequence 5098, App
29	294.6	31.1	824	9	US-09-988-442-75	Sequence 75, Appl
30	294.6	31.1	824	9	US-10-073-865-64	Sequence 64, Appl
31	294.6	31.1	824	9	US-10-103-313-230	Sequence 230, App
32	294.6	31.1	824	10	US-09-764-853-351	Sequence 351, App
33	291.6	30.8	2282	9	US-10-098-841-92	Sequence 92, Appl
34	291.6	30.8	2282	10	US-09-747-835A-62	Sequence 62, Appl
35	290	30.6	1788	10	US-09-747-835A-35	Sequence 35, Appl
36	290	30.6	2735	10	US-09-747-835A-34	Sequence 34, Appl
37	285.2	30.1	916	9	US-09-974-591-15	Sequence 15, Appl
38	284.2	30.0	1957	10	US-09-864-761-13963	Sequence 13963, A
39	282.8	29.8	933	10	US-09-864-761-30527	Sequence 30527, A
40	276.4	29.2	939	10	US-09-886-055-66	Sequence 66, Appl
41	276.4	29.2	1050	9	US-09-974-591-1	Sequence 1, Appli
42	276.4	29.2	1050	9	US-09-974-591-3	Sequence 3, Appli
43	276.4	29.2	1050	9	US-09-974-591-7	Sequence 7, Appli
44	275.4	29.1	939	10	US-09-886-055-440	Sequence 440, App
45	274.8	29.0	1050	9	US-09-974-591-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-10-023-601-7  
: Sequence 7, Application US/10023601  
: Publication No. US20030083463A1  
: GENERAL INFORMATION:  
: APPLICANT: Padigaru, Muralidhara  
: APPLICANT: Kekuda, Ramesh  
: APPLICANT: Colman, Steven  
: APPLICANT: Spytek, Kimberly  
: APPLICANT: Ballinger, Robert  
: APPLICANT: Vernet, Corine  
: APPLICANT: Li, Li  
: APPLICANT: Shenoy, Suresh  
: APPLICANT: Casman, Stacie  
: APPLICANT: Guzev, Vladimir  
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
: FILE REFERENCE: 21402-224AF  
: CURRENT APPLICATION NUMBER: US/10/023,601  
: PRIOR FILING DATE: 2001-12-18  
: PRIOR APPLICATION NUMBER: 60/256,635  
: PRIOR FILING DATE: 2000-12-18  
: PRIOR APPLICATION NUMBER: 60/259,743  
: PRIOR FILING DATE: 2001-01-04  
: PRIOR APPLICATION NUMBER: 60/299,327  
: PRIOR FILING DATE: 2001-06-19  
: PRIOR APPLICATION NUMBER: 60/261,498  
: PRIOR FILING DATE: 2001-01-12  
: PRIOR APPLICATION NUMBER: 60/263,689  
: PRIOR FILING DATE: 2001-01-24  
: PRIOR APPLICATION NUMBER: 60/267,464  
: PRIOR FILING DATE: 2001-02-08  
: PRIOR APPLICATION NUMBER: 60/271,021  
: PRIOR FILING DATE: 2001-02-22  
: PRIOR APPLICATION NUMBER: 60/275,946  
: PRIOR FILING DATE: 2001-03-14  
: PRIOR APPLICATION NUMBER: 60/278,150  
: PRIOR FILING DATE: 2001-03-23  
: PRIOR APPLICATION NUMBER: 60/285,718  
: PRIOR FILING DATE: 2001-04-23  
: PRIOR APPLICATION NUMBER: 60/312,902  
: PRIOR FILING DATE: 2001-08-16  
: PRIOR APPLICATION NUMBER: 60/257,876

```
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/260,718
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/284,591
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1172
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-601-7
```

Query Match 98.7%; Score 935.4; DB 9; Length 1172;

Best Local Similarity 99.8%; Pred. No. 1.9e-287;

Matches 947; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 ATGGGAAGATGGTGAACCAAGTCCCTACACAGATGGCTTCTTCTTGGGCATCTTTCC 60
    |||
DB 3 ATGGGAAGATGGTGAACCAAGTCCCTACACAGATGGCTTCTTCTTGGGCATCTTTCC 62

QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGACATTATGGTGTCTTACAGTGGCCCTC 120
    |||
DB 63 CACAGCCAGACTGACCTTGTCTCTCTCTCTGACATTATGGTGTCTTACAGTGGCCCTC 122

QY 121 TGTGGGAATGTCCTCTCATCTTCTCATCTACCTGACGCTGACCTTCACACCCCATG 180
    |||
DB 123 TGTGGGAATGTCCTCTCATCTTCTCATCTACCTGACGCTGACCTTCACACCCCATG 182

QY 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTGTGTCTGTACATTGTGCCA 240
    |||
DB 183 TACTTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTGTGTCTGTACATTGTGCCA 242

QY 241 AAGATGGCAGCCCAACTTCTGTCTGAGGAGGAGTCCATCTCTTGTGGGCTGTGCCATA 300
    |||
DB 243 AAGATGGCAGCCCAACTTCTGTCTGAGGAGGAGTCCATCTCTTGTGGGCTGTGCCATA 302

QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCT 360
    |||
DB 303 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCT 362

QY 361 TATGACCACTAGCTGGCCGTAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 420
    |||
DB 363 TATGACCCGTAGCTGGCCGTAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 422

QY 421 GTCTGTCTCCAGATTAAGTGGAGCTCTGGCCCTTTGGGATTAATAGATGAGTGAATCAG 480
    |||
DB 423 GTCTGTCTCCAGATTAAGTGGAGCTCTGGCCCTTTGGGATTAATAGATGAGTGAATCAG 482.

QY 481 ATGTGGCAGCCATGGGCTTACCTTAAGTGGCTCAAGGAGCGTGATCACTTTTCTGT 540
    |||
DB 483 ATGTGGCAGCCATGGGCTTACCTTAAGTGGCTCAAGGAGCGTGATCACTTTTCTGT 542

QY 541 GAGTACCAAGCTTTATGAAGCTGGGCTGTGACAGACCTTCCCTTTTGACACCCCTCTC 600
    |||
DB 543 GAGTACCAAGCTTTATGAAGCTGGGCTGTGACAGACCTTCCCTTTTGACACCCCTCTC 602

QY 601 TTTGCTGTCTGTCTCATGCTTCTCTTCCCTCTCCATCATGAGCCCTCTATGCT 660
    |||
DB 603 TTTGCTGTCTGTCTCATGCTTCTCTTCCCTCTCCATCATGAGCCCTCTATGCT 662

QY 661 TGCATC-CTAGGGGCTGTGCTCCGATACGCTGTGCTCAGGCGTGAAGGAGCCCTGCG 719
    |||
DB 663 TGCATCTTAGGGGCTGTGCTCCGATACGCTGTGCTCAGGCGTGAAGGAGCCCTGCG 722

QY 720 CACCTGTCTCTCCACCTAACAGAGTGTACACCTCTTCTATGGGGCAGCCATGTTCATGTA 779
    |||
DB 723 CACCTGTCTCTCCACCTAACAGAGTGTACACCTCTTCTATGGGGCAGCCATGTTCATGTA 782

QY 780 CCTGAGGCGCTAGGCGCTACCGGGCGCTAGCCATGACAAAGGTGCGCTCTATCTTACAC 839
    |||
DB 783 CCTGAGGCGCTAGGCGCTACCGGGCGCTAGCCATGACAAAGGTGCGCTCTATCTTACAC 842
```

```
QY 840 AGTCCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGAGG 899
    |||
DB 843 AGTCCTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGAGG 902

QY 900 GGCAGTGAAGAAGGGCTGACCGCTGACAGATTTGGCAGCCAGCACTGA 948
    |||
DB 903 GGCAGTGAAGAAGGGCTGACCGCTGACAGATTTGGCAGCCAGCACTGA 951
```

## RESULT 2

US-09-886-055-176

; Sequence 176, Application US/09886055

; Patent No. US20020132273A1

; GENERAL INFORMATION:

; APPLICANT: STRYER, LUBERT

; APPLICANT: ZOZULYA, SERGEY

; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND

; FILE REFERENCE: 078003-0277150

; CURRENT APPLICATION NUMBER: US/09/886,055

; CURRENT FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,812

; PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 522

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 176

; LENGTH: 948

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-886-055-176

Query Match 85.5%; Score 810.4; DB 10; Length 948;

Best Local Similarity 90.9%; Pred. No. 1.1e-247;

Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
QY 1 ATGGGAAGATGGTGAACCAAGTCCCTACACAGATGGCTTCTTCTTGGGCATCTTTCC 60
    |||
DB 1 ATGGGAAGATGGTGAACCAAGTCCCTACACAGATGGCTTCTTCTTGGGCATCTTTCC 60

QY 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGACATTATGGTGTCTTACAGTGGCCCTC 120
    |||
DB 61 CACAGCCAGACTGACCTTGTCTCTCTCTCTGACATTATGGTGTCTTACAGTGGCCCTC 120

QY 121 TGTGGGAATGTCCTCTCATCTTCTCATCTACCTGACGCTGACCTTCACACCCCATG 180
    |||
DB 121 TGTGGGAATGTCCTCTCATCTTCTCATCTACCTGACGCTGACCTTCACACCCCATG 180

QY 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTGTGTCTGTACATGTGCCA 240
    |||
DB 181 TACTTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTGTGTCTGTACATGTGCCA 240

QY 241 AAGATGGCAGCCCAACTTCTGTCTGAGGAGGAGTCCATCTCTTGTGGGCTGTGCCATA 300
    |||
DB 241 AAGATGGCAGCCCAACTTCTGTCTGAGGAGGAGTCCATCTCTTGTGGGCTGTGCCATA 300

QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCT 360
    |||
DB 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCT 360

QY 361 TATGACCACTAGCTGGCCGTAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 420
    |||
DB 361 TATGACCCGTATGAGCCATGAGCCACCACCTTCACTATCCCATCTCATGATCAGAGG 420

QY 421 GTCTGTCTCCAGATTAAGTGGAGCTCTGGCCCTTTGGGATTAATAGATGAGTGAATCAG 480
    |||
DB 421 GTCTGTCTCCAGATTAAGTGGAGCTCTGGCCCTTTGGGATTAATAGATGAGTGAATCAG 480

QY 481 ATGTGGCAGCCATGGGCTTACCTTAAGTGGCTCAAGGAGCGTGATCACTTTTCTGT 540
    |||
DB 481 ATGTGGTAGTAATGAATTTCCCTTACTGTGCTGAGGAGGTGAACCATTTCTTGT 540

QY 541 GAGTACCAAGCTTTATGAAGCTGGGCTGTGACAGACCTTCCCTTTTGACACCCCTCTC 600
    |||
```

```
Db 541 GAGATGCTATCCTTGTGAAGCTGGCCCTGTGTAGACATCCCTTTTGAGAAGTGATA 600
QY 601 TTTGCTGCTGTGTCTTCATGCTTCCTTCCTTCCTTCATCATCATGAGCCCTCATGCT 660
    |||||
Db 601 TTTGCTGCTGTGTCTTCATGCTTCCTTCCTTCCTTCATCATCATGAGCCCTCATGCT 660
QY 661 TGCATCTAGAGGGGTGTGCTCCGAATAGCTCTGTCTAGCCCTGGAATAAGCCCTGGCC 720
    ||| ||||| ||||| ||| |||||
Db 661 CACATTTCTAGGAGAGTGTGCTGCAATGCACTCTGCTCAGGCCCTGGAATAAGCCCTGGCC 720
QY 721 ACCGCTCTCCCTCCACCTAACAGCTGTACCCCTCTCTCTATGAGGCCATGTTCATGTAC 780
    |||||
Db 721 ACCGCTCTCCCTCCACCTGACAGCTGTACCCCTCTCTCTATGAGGCCATGTTCATGTAC 780
QY 781 CTGAGGCTTAGGCGCTACCGGGCCCTAGCCATGACAAGTGGCCCTCTATCTTCTACACA 840
    |||||
Db 781 CTGAGGCTTAGGCGCTACCGGGCCCTAGCCATGACAAGTGGCCCTCTATCTTCTACACG 840
QY 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGAATGGGAGGTGATGGGG 900
    |||||
Db 841 GTCTTACTTCCATGCTGAACCCCTCATTTACAGCTTGAGAATGGGAGGTGATGGGG 900
QY 901 GCACGTAGGAAGAGGGGTGAGCCGCTGCAGGATTTGGCAGCCAGCACTGA 948
    |||||
Db 901 GCACGTAGGAAGAGGGGTGAGCCGCTGCAGGATTTGGCAGCCAGCACTGA 948
```

RESULT 3  
US-10-259-430-19

```
; Sequence 19, Application US/10259430
; Publication No. US20030082615A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Randall
; APPLICANT: Yau, King-wai
; APPLICANT: Krautwurst, Dietmar
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries
; TITLE OF INVENTION: ad Methods of Making and Using Them
; FILE REFERENCE: 001107.00105
; CURRENT APPLICATION NUMBER: US/10/259,430
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/465,901
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/112,605
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: PCR primer
; US-10-259-430-19
```

Query Match 55.5%; Score 526.6; DB 9; Length 669;  
Best Local Similarity 86.7%; Pred. No. 2.1e-157;  
Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
QY 193 AGCCAGCTCTCCCTCATGACCTCATGTGTGTGTATACATTTGCCCAAGATGGCAGCC 252
    |||||
Db 1 AGTCAGCTCTCCCTCATGACCTCATGTGTGTGTATACATTTGCCCAAGATGGCAGTC 60
QY 253 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATACAAATGGCTTT 312
    |||||
Db 61 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATACAAATGGCATTT 120
QY 313 TTTGCTCTCTTTGGGATCTGAGGGGCTCTTGTGGGACTCATGGCTTATGACCACCTAC 372
    |||||
Db 121 TTTGCTCTCTTTGGGATCTGAGGGGCTCTTGTAGGACTCATGGCTTATGATGCTAT 180
QY 373 GTGGCCGTAGCCACCACCTTCACTATCCCATCTCATGAATCAGAGGGTGTGTCCAG 432
    |||||
Db 181 GTGGCCATTAGCCACCACCTTCACTATCCCATCTCATGAAGCCAAAGGTGTGTCTCCAG 240
```

```
QY 433 ATTACTGGGAGCTCCTGGGCTTTGGGATATATAGATGAGTGAATGATGATGGCAGCC 492
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ATTGCTGGAAGTTCTCTGGGCTTTGGGATCCTTGATGGAATTAATTCAGATGTGGCAGCC 300
QY 493 ATGGCTTACCTTACTGTGGCTCAAGAGCGGTGGATCACTTTTCTGTGAGGTAGAGCT 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ATGAGCTGCCCTTACTGTGGCTCAGCGTATATAGATCACTTCTTGTGAAGTGGCGCT 360
QY 553 TTATTTGAAGCTGGCTGTGACAGACTTCCCTTTTGGACACCCCTCTTTGCTTCTGT 612
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TTACTGAAAGCTGGCTGTGACAGACTTCCCTTTTGGACACCCCTCTTTGCTTCTGT 420
QY 613 GTCTTCATGCTTCTCTCTCCCTTCTTCATCATGATGAGCCCTCTATGCTTGCATGAGG 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GTCTTATGCTGTCTCTCTCTCTTCTTCGATCATGTGACTTCTATGCTCGCATCTGGGG 480
QY 673 GCTGTGCTCCGAATACGCTCTGCTCAGGCGCTGGAATAAGCCCTGACCACCTGCTCTCC 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 GCTGTGCTCCGATACACTCTGCCCCAGTCCGAAAAAAGGCCCTGACCACCTGCTCTCC 540
QY 733 CACCTAACAGCTGTACACCCCTCTTCATGAGGGGACGCAATGTTACATGATGAGCCCTAGG 792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CACCTAACAGCTGTCTCTCTCTTCATGAGGGGACGCAATGTTACATGATGAGCCCAAG 600
QY 793 CGCTACCGGGCCCTAGCCATGACAGGTGGCTCTATCTTCTACAGTCTTACTCTCC 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 CGATATCGCGCTCTAGCCATGACAGAAAGTTGTCTCAATCTTCTACAGTCTTACTCTCC 660
QY 853 ATGCTGAAC 861
    ||||| |||
Db 661 ATGCTCAAC 669
```

RESULT 4  
US-10-259-430-27

```
; Sequence 27, Application US/10259430
; Publication No. US20030082615A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Randall
; APPLICANT: Yau, King-wai
; APPLICANT: Krautwurst, Dietmar
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries
; TITLE OF INVENTION: ad Methods of Making and Using Them
; FILE REFERENCE: 001107.00105
; CURRENT APPLICATION NUMBER: US/10/259,430
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/465,901
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/112,605
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: PCR primer
; US-10-259-430-27
```

Query Match 55.5%; Score 526.6; DB 9; Length 669;  
Best Local Similarity 86.7%; Pred. No. 2.1e-157;  
Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
QY 193 AGCCAGCTCTCCCTCATGACCTCATGTGTGTGTATACATTTGCCCAAGATGGCAGCC 252
    |||||
Db 1 AGTCAGCTCTCCCTCATGACCTCATGTGTGTGTATACATTTGCCCAAGATGGCAGTC 60
QY 253 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATACAAATGGCTTT 312
    |||||
Db 61 AACTTCCTGTCTGGCAGGAAGTCCATCTCTTTGTGGGCTGTGGCATACAAATGGCATTT 120
```

```
QY 313 TTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGACCACTAC 372
    |||
Db 121 TTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGACCACTAC 180
QY 373 GTGGCGCTTAGCCACCCACTTCACATATCCCATCCATGATGATGAGGGTCTGTCCAG 432
    |||
Db 181 GTGGCGCTTAGCCACCCACTTCACATATCCCATCCATGATGATGAGGGTCTGTCCAG 240
QY 433 ATTACTGGAGCTCCTGGGCTTTGGGATATAGATGAGTATGATGATGAGTGGAGCC 492
    |||
Db 241 ATTGCTGGAAGTCTCTGGGCTTTGGGATCCTGATGGAATATATGATGATGAGTGGAGCC 300
QY 493 ATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGATCACTTTTCTGTGAGTACAGCT 552
    |||
Db 301 ATGAGCGTGCCTTACTGTGGCTCACGGTATATAGATCACTTCTGTGAAGTGGCGCT 360
QY 553 TTATGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGT 612
    |||
Db 361 TTACTGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGT 420
QY 613 GTCTTCATGCTTCTCTCCCTTCCCATCATCATGAGGCTCTATGCTTGCATCCTAGG 672
    |||
Db 421 GTCTTCATGCTTCTCTCCCTTCCCATCATCATGAGGCTCTATGCTTGCATCCTAGG 480
QY 673 GCTGTGCTCCGAATACGCTCTGTCTCAGAGGCTTGAAAAAGCCCTGGCCACCTGCTCTCC 732
    |||
Db 481 GCTGTGCTCCGAATACGCTCTGTCTCAGAGGCTTGAAAAAGCCCTGGCCACCTGCTCTCC 540
QY 733 CACCTAACAGCTGTACACCTCTTCTATGGGGGAGCCATGTTCATGTACCTGAGGCTTAGG 792
    |||
Db 541 CACCTAACAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 793 CGCTACCGGGGCCCCCTAGCCATGACAGGTTGGCTCTATCTTCTACACAGTCTTACTCCC 852
    |||
Db 601 CGATATCGCGCTCCTAGCCATGACAAAGTTGTCTCAATCTTCTACACAGTCTTACTCCT 660
QY 853 ATGCTGAAC 861
    |||
Db 661 ATGCTCAAC 669

RESULT 5
US-10-259-430-23
; Sequence 23, Application US/10259430
; Publication No. US20030082615A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Randall
; APPLICANT: Yau, King-wai
; APPLICANT: Krautwurst, Dietmar
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries
; TITLE OF INVENTION: ad Methods of Making and Using Them
; FILE REFERENCE: 001107.00105
; CURRENT APPLICATION NUMBER: US/10/259,430
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/465,901
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/112,605
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: PCR primer
US-10-259-430-23

Query Match 55.48; Score 525; DB 9; Length 669;
Best Local Similarity 86.58; Pred. No. 6,7e-157;
Matches 579; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
```

QY 193 AGCCAGCTCTCCCTCATGAGCCTCATGTGTGTGTAACATTGTGCCAAAGATGGCAGCC 252

```
Db 1 AGTCACTCTCCCTCATGAGCCTCATGTGTGTGTAACATTGTGCCAAAGATGGCAGTC 60
QY 253 AACTTCCTGTCTGGGAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATTGGCTT 312
    |||
Db 61 AACTTCCTGTCTGGGAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATTGGCTT 120
QY 313 TTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGACCACTAC 372
    |||
Db 121 TTGTCCTCTTGTGGGATCTGAGGGGCTTGTGGGACTCATGGCTTATGATGCTAT 180
QY 373 GTGGCGCTTAGCCACCCACTTCACATATCCCATCCATGATGATGAGGGTCTGTCCAG 432
    |||
Db 181 GTGGCGCTTAGCCACCCACTTCACATATCCCATCCATGATGATGAGGGTCTGTCCAG 240
QY 433 ATTACTGGAGCTCCTGGGCTTTGGGATATATAGATGAGTATGATGATGAGTGGAGCC 492
    |||
Db 241 ATTGCTGGAAGTCTCTGGGCTTTGGGATCCTTGTGAAGTATGATGATGAGTGGAGCC 300
QY 493 ATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGATCACTTTTCTGTGAGTACAGCT 552
    |||
Db 301 ATGAGCGTGCCTTACTGTGGCTCACGGTATATAGATCACTTCTGTGAAGTGGCGCT 360
QY 553 TTATGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGT 612
    |||
Db 361 TTACTGAAGCTGGGCTGTGACAGACACTCCCTTTTGACACCCCTCTTGTGCTGT 420
QY 613 GTCTTCATGCTTCTCTCCCTTCCCATCATCATGAGGCTCTATGCTTGCATCCTAGG 672
    |||
Db 421 GTCTTCATGCTTCTCTCCCTTCCCATCATCATGAGGCTCTATGCTTGCATCCTAGG 480
QY 673 GCTGTGCTCCGAATACGCTCTGTCTCAGAGGCTTGAAAAAGCCCTGGCCACCTGCTCTCC 732
    |||
Db 481 ACTGTGCTCCGATACACTCTGTGCCAGTCCGAAAAAGCCCTGGCCACTTGTCTCTCC 540
QY 733 CACCTAACAGCTGTACACCTCTTCTATGGGGGAGCCATGTTCATGTACCTGAGGCTTAGG 792
    |||
Db 541 CACCTAACAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 793 CGCTACCGGGGCCCCCTAGCCATGACAGGTTGGCTCTATCTTCTACACAGTCTTACTCCC 852
    |||
Db 601 CGATATCGCGCTCCTAGCCATGACAAAGTTGTCTCAATCTTCTACACAGTCTTACTCCT 660
QY 853 ATGCTGAAC 861
    |||
Db 661 ATGCTCAAC 669
```

```
RESULT 6
US-09-747-155-212
; Sequence 212, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
; APPLICANT: Rouquier, Sylvie
; APPLICANT: Giorgi, Dominique
; TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding
; FILE REFERENCE: 19904-008 (C009B6834US)
; CURRENT APPLICATION NUMBER: US/09/747,155
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/171,746
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 431
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 212
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(488)
; OTHER INFORMATION: Taxon = 9593; gene = GGO101; pseudogene; Accession DDBJ/EMBL/G
US-09-747-155-212
```

Query Match 41.2%; Score 390.2; DB 10; Length 488;  
Best Local Similarity 88.9%; Pred. No. 4.8e-114;  
Matches 433; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

```
QY 373 GTGGCGGTAGCCACCCACTTCACTATCCCATCTCATGAATCAGAGGCTGTCTCCAG 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GTGGCCATTAGCCACCCACTTCACTATCCCATCTCATGAATCAGAGGCTGTCTCCAG 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 433 ATTAATGGAGCTCCTGGGCTTTGGGATTAATAGATGG-AGTGAATTCAGATGGTGCCAGC 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 ATTAACGGGAGCTCCTGGGCTTTGGGATTAATCAGATGGCTTTGATCCAGATGGTGAGT 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 492 CATGGGCTTACCTACTGTGGCTCAAGAGCGTGATCATTCTTTCTGTGAGGTACAGC 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 AATGAATTTCCCTACTGTGGCTGAGGAGGTAACCATTTCTTCTGTGAGATGCTATC 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 552 TTATTTGAAGCTGGGCTGTGACACACTTCCCTTTTGACACCCCTCTTGTCTGCTG 611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 CTGTGTGAAGCTGGGCTGTGACACACTTCCCTTTTGAGAGGTGATATTTGCTTGTG 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 612 TGTTCTCATGCTTCTCTCTTCCCTTCCATCATATGAGGCTCTCTATGCTTGCATCTAGG 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TGTTCTCATGCTTCTCTCTTCCCTTCCATCATATGAGGCTCTCTATGCTTGCATCTAGG 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 672 GCGTGTGCTCCGAATACGCTCTGCTCAGGCGCTGGAAGAAAGCCCTGGCCACCTGCTCTC 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 GACTGTGCTGCAATATGCACTCTGCTCAGGCGCTGGAAGAAAGCCCTGGCCACCTGCTCTC 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 732 CCACCTAAGCTGTGACCCCTCTTCTATGAGGCGACCCATGTTCAATGTACCTGAGGCTTAG 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 CCACCTGACAGCTGTGACCCCTCTTCTATGAGGCGACCCATGTTCAATGTACCTGAGGCTTAG 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 792 GCGCTACCGGGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACACAGTCCCTACTCC 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 GCGCTACCGGGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACACAGTCCCTACTCC 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 852 CATGCTG 858
    ||||| |||||
Db 482 CATGCTG 488
```

RESULT 7  
US-09-826-508-23  
; Sequence 23, Application us/09826508  
; Patent No. US20010025099A1  
; GENERAL INFORMATION:  
; APPLICANT: Nabil Elshourbasy  
; APPLICANT: Lisa Vawter  
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides  
; TITLE OF INVENTION: and Polynucleotides  
; FILE REFERENCE: GP-70744USB  
; CURRENT APPLICATION NUMBER: US/09/826,508  
; CURRENT FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 23  
; LENGTH: 1290  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
; US-09-826-508-23

Query Match 41.0%; Score 388.8; DB 10; Length 1290;  
Best Local Similarity 64.4%; Pred. No. 2.1e-113;  
Matches 582; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

```
QY 129 TGTCCTCTCATCTTCTCATCTACCTGAGCGTGGACTTCACACCCCATGTACTTCTT 188
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 CTCTGTATGCTGTCTCTCTCATCTACCTGAGACACCCAGCTCCACACCCCATGTACTTCTT 480
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 189 CCTAGCCAGCTCTCCCTCATGTGACCTCATGTGTCTGTGAACATTTGTGCCAAGATGGC 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CCTAGCCAGCTCTCCCTCATGTGACCTCATGTGTCTGTGAACATTTGTGCCAAGATGGC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 249 AGCCAACTTCTCTGTGGCAGAGATCCATCTCCCTTGTGGGCTGTGGCATACAAATTTGG 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CTTCAACTACCTGTCTGTGGCAGAGATCCATTTCTATGGCTGTGTGCCACACAATTTT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 309 CTTTTTGTCTCTCTGTGGATCTGAGGGCTCTTGTCTGGGCTGTGGCATACAAATTTGG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 CTTCATATACATCACTGCTTGGCTCTGAATGCTTCTTGTGGCTGTGTATGAGCCG 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 369 CTACGTGGGCTTACCCACCCACTTCACTATCCCATCTCATGAATCAGAGGCTGTCT 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 CTACACTGCCATTTGCCACCCCTTAAGATACACCAATCTCATGAGCCCTAAATTTGTGG 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 429 CCAGATTTACTGGGACCTCTCTGGGCTTTGGGATTAATAGATGAGATGATTCAGATGCTGGC 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 ACTTATGACTGCTCTTCTCTGATCTGAGGCTCTACAGATGAGATCATTTATGCTGTAGC 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 489 AGCCATGGGCTTACCTTACTGTGCTCAAGAGCGTGATCATTCTTCTGTGAGGTACA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 CACATTTTCTCTCTCTACTGTGAGGCTCTCGGGAATAGCCACCTTCTGTGAGGTACC 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 549 AGCTTATTTGAAGCTGGGCTGTGACACACTTCCCTTTTGACACCCCTCTTGTGCTTG 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 TTCCCTACTAATCTCTCATGCAATGACACATCAATATTTGAAAGGTATTTTCAATTTG 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 609 CTGTGCTTCTCATGCTTCTCTTCCCTTCTCCATCATATGAGGCTCTTATGCTTGCATCT 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 CTCTATAGTAATGCTTGTCTTCCCTGTGCAATCATATGCTTCTTATGCTGAGGTAT 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 669 AGGGCTGTGCTCCGAATACGCTCTGCTCAGGCGCTGGAAGAAAGCCCTGGCCACCTGCTC 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 TCTGCTGTATCTCATGATGGATCTGAGAGGGCTCGCAAGCTTTTACAGACCTGTT 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 729 CTCCACCTAAGCTGTACCCCTTCTATGAGGCGACCCATGTTCAATGTACCTGAGGCC 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 CTCTCACTCATGCTGTGGAGATTTCTATGAGCAGGTTTGTTCATGTACATACAGCC 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 789 TAGGCGCTACCGGGCCCTAGCCATGACAGAGTGCGCTCTATCTTCTACACAGTCTTAC 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 CACATCTGATCGCTCCCAAGCGCAGACAAAGCTGTGTCTGTATTTACACCATCTCTAC 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 849 TCCCATGCTGAACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGGGCAGCTGAG 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 TCCCATGCTGAATCCCTCATCTACAGCCCTCCGCAACAAGAGTACAGAGCATTCAT 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 909 GAAG 912
    |||||
Db 1201 GAAG 1204
```

RESULT 8  
US-09-886-055-430  
; Sequence 430, Application us/09886055  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886,055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213,812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1



; SEQ ID NO 430  
; LENGTH: 975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-886-055-430

Query Match 38.2%; Score 361.8; DB 10; Length 975;  
Best Local Similarity 62.6%; Pred. No. 7.2e-105;  
Matches 564; Conservative 0; Mismatches 337; Indels 0; Gaps 0;

```
QY 26 ACACAGATGCTCTCTCTCTGCGCATCTTCCACAGCCAGACTGACCTGTGCTCT 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 ACTCCACTACTCTGCTCTCACAGGCTCATACCCATCTGCTTCCCGGCTCTCT 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 86 TCTCTGAGTTATGGTGTCTTCAAGTGGCCCTCTGTGGAAATGCTCTCATCTTCC 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 TTGCAATAGTCTTCTCCATCTTGTGGTGGCTATACAGCCAACTGGTCATGATCTTCC 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 146 TCATCTACCTGGAGCTGGACTTCAACCCCATGACTTCTTCCAGCCAGCTTCC 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 TCATCCACATGAGATCCCGCTCCACACACCAGTACTTCTTGTCTCAGCCAGCTTCCA 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 206 TCATGGACCTCATCTGTGTCTGTAAATTTGCAAGATGGCAAGCACTTCTGTCTG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 TCATGGATACCATCTACATCTGTATCACTGTCCCAAGATGTCCAGACCTCCGTCCA 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 266 GCAGGAAGTCCATCTCTTGTGGGCTGTGCATACAATTTGGCTTTTGTCTCTCTG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 AGGACAGAGACCATTTCTCTGCTGCTGTGCAGTTCAGATCTTCTTACCTGACCTTGA 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 326 TGGATCTGAGGGCTCTTGTGAGGACTCATGGCTTATGACCAGTGGCGCTTACCC 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 TTGAGAGGGGAATTTCTCTGCTGGGTCTCATGGCTTATGACCGCTATGTGGCTGTGCA 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 386 ACCCACTTCACTATCCATCTCTCAATGAATCAGAGGCTCTCTCAGATTACTGGAGCT 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ACCCTCTACGTAACCTCTCTCTCAATGAACCGAGGCTTGTCTTATCATGTGTGGCT 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 446 CCTGGGCTTTGGGATATAGATGAGTATGATGAGTGGCAAGCATGGGCTTACCTT 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 CCTGGGCTTTGGGATATAGATGAGTATGATGAGTGGCAAGCATGGGCTTACCTT 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 506 ACTGTGGCTCAAGAGCGTGAATCACTTTTCTGTGAGGTACAAAGCTTATGAAGCTGG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 TCTGTAGATCCCGAGAGATCAATCACTTTTCTGTGAGATCCCAAGCGCTGTGAAGTTGT 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 566 CCTGTGCAGACACTTCCCTTTTGAACCCCTCTCTTGTGTGTGTCTTCAATGCTTTC 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 CTTGCACAGACACTCTCTATGAGACCTGTATGCTGTGTGTGTGTGTGTGTGTGTGT 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 626 TCTCTCCCTTCTCATCATCATGAGCTCTATGCTTGCATCCTAGGGGCTGTGTCCGAA 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 TTTATCCCTCTATCTGTCTCTGTCTCTACAGGACATCTCTGACTGTCCACAGGA 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 686 TAGGCTCTGCTCAGGCTTGAAGAAAGCCCTGGCAAGCTGCTCTCCACCTAACAGCTG 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 TGAAGCTCTGCTGAGGGCCGCGCAAGAGCTTGTCTAGCTGTCTCCACATATATGTGG 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 746 TCACCTCTTCTATGAGGCGAGCATGTATGATGAGCTAGGCTAGGCGCTACCGGGCC 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 TGAGGCTTTTCTACGGGGCAGCCTTCTACACCAAGTGTGCTCCACTCTACCAACTC 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 806 CTAGCATGACAAGTGGCTCTATCTTCTACACAGTCTTACTCCATGCTGAACCCCC 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 CAGAGAAAGATAAAGTGTGTCTGCTTCTACACCAATCTTCAACCCCATGCTCAACCCAC 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 866 TCATTTACAGCTTGAAGATGGGAGTGTATGGGGCACTGAGGAAGGGCTGACCGCT 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 TCATTTACAGCTTGAAGATTAAGATGTGGCTGACAGCTCTGAGGAAGTACTAGGGAGAT 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 926 G 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 G 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9  
US-09-886-055-478  
; Sequence 478, Application US/09886055  
; Patent No. US20020132273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886, 055  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 478  
; LENGTH: 951  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-886-055-478

Query Match 38.0%; Score 360.6; DB 10; Length 951;  
Best Local Similarity 62.7%; Pred. No. 1.7e-104;  
Matches 561; Conservative 0; Mismatches 334; Indels 0; Gaps 0;

```
QY 33 TGGCTTCTCTCTTGGGATCTTTCCACAGCCAGACTGACCTGTCTCTCTCTG 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 TGACTTCAACCTCTCTGCGGCTTGTGTGAACAGTGAAGGCTGCGGATTTGATTAACAGT 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 93 AGTTATGTGTGTTCAACAGTGGCCCTCTGTGGAAATGCTCTCATCTTCTCATCTA 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 GATCCTTGTCTGTCTTCTTGGGCGCTGACTGCAAAATTTGTATGATATTCTTGAATCA 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 153 CCTGAGCGCTGAGCTTACACACCCCATGTAATCTTCTCTCAGCCAGCTCTCTCATGGA 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 GGTGAGACTCTGCTTCCACACCCCATGTAATCTTCTCTCAGCCAGCTCTCTCATGGA 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 213 CCTCATGTGTGTTCAACATTTGCCAAGATGGCAGCCAACTTCTCTGTGCGAGGAA 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CACCTTTTCACTGTATCCACTGTCCCAAAACTCTGCGAGACATGTTCTTAAGAGAA 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 273 GTCCATCTCTTGTGGGCTGTGGCATACAAATTTGCTTTTGTCTCTCTGTGGGATC 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 GATCATTTCTCTTGTGGGCTGTGGCATACCAATTTCTCTCTCATGATGATGTTG 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 333 TGAGGGGCTCTTGTGGGACTCATAGGCTTATGACCACTACGTGGCGCTTAGCCACCT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 TGAGTTCTTCTCTGTGGGCTCATAGGCTTATGACCTACGTGGCTGTCTGAACCTCT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 393 TCATATCCCATCTCATGAATGAAGAGGCTGTGCTCCAGATTTACTGGAGCTCTGGGC 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 GAGATACCCAGTCTGTATGAACCCCAAGAGTGTCTTTGTGTGTGTGTGTGTGTGT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 453 CTTTGGGATATAGATGAGTGTATGAGTGTGGGAGCCATGGGCTTACCTTACTGTGG 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 TGGGGGCTCCCTGATGGCTTGTGCTCACTCCCATCACCATGAATGTCCCTTACTGTGG 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 513 CTCAAGAGCGTGTACATCTTTTCTGTGAGGTACAAAGCTTATGAAGCTGGCTGTGC 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 CTCCGAAGTATCAACATTTTCTGTGAGATCCAGAGTGTGAAGTGTGCTGTGC 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 573 AGACACTTCCCTTTTGAACACCCTCTTGTGCTGTGTCTTCAATGCTTCTCTTCC 632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 AGACAGTCTCTGTATGAACCTGTATGATGATGATGATGATGATGATGATGATGATGAT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 633 CTTTCCATCATCATGAGCTCTCTATGCTTGTGATCTAGGGGCTGTGCTCCGAATAGCTC 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 CATCTCTATCATCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 693 TGCTAGGCGCTGAAAAAAGCCCTGGCCACCTGCTCTCCACCTAACAGCTGTACCCCT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

Db 681 TGCTGAGGTGCGAAAAAGGCCCTTACCACCTGTGTTCCCTCCACTTGACTGTAGTTAGCAT 740
QY 753 CTTCATATGGGGGACGACATGTTCATGTACCTGAGGCGCTAGGCGCTACCGGGCCCTAGGCCA 812
Db 741 CTTCATATGGGGGCTGCTTCTACACATAGGTGCTGCCCCAGTCCCTTCCACACCCCCGAGCA 800
QY 813 TGACAGGTGGCCCTATCTCTACACAGTCCCTTACTCCCATGCTGAACCCCTCATTTA 872
Db 801 GGACAAAGTAGTGTACGCCCTTCTATACCATTTGTCACGCCCATGCTTAATCTCTCATCTA 860
QY 873 CAGCTTGAGGAATGGGAGAGGTGATGGGGGCGACACTGAGGAAGGGGCTGGACCGCTGC 927
Db 861 CAGCCTCAGAAACAGACGCTCATAGGGGCAATTAAGGTAATTTGCATGTTGC 915
```

RESULT 10

```

US-09-886-055-436
; Sequence 436, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 436
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-436
```

Query Match 36.4%; Score 345; DB 10; Length 948;  
Best Local Similarity 61.3%; Pred. No. 1.6e-99;  
Matches 555; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

```

QY 7 AGATGGGTGAGCAGTCTACACAGATGGCTTCTCTCTTGGGCATCTTTCCACAGC 66
Db 16 AGATGGGCCACACACTGGAAGTTGGATTTCATCTCATGAGCACTCTTACAGCATCC 75
QY 67 CAGACTGACCTTGTCTCTCTCTGCAAGTTAGTGTGTTCTTACAGTGGCCCTCTGTGG 126
Db 76 AACATCCAGCTCTACTAGTGTGTCATCTTGTGTTTCTTGAAGGCGTGTCTGGA 135
QY 127 AATGCTCTCTCTCTCTCTCTCATCTACCTGAGCGCTGACTTTCACACCCCATGTACTTC 186
Db 136 AATGCTGTCTCTCTCTCTCTCATGACATGGCGCTACATTTCTGTCACTGTGCCAAGATG 195
QY 187 TTCCTCAGCCAGCTCTCCCTCATGAGACCTCATGTTGTTGTTAATGTTGTCGAAGATG 246
Db 196 TTCATCAGTCAATTTGTTCTCTCATGACATGGCGTACATTTCTGTCACTGTGCCAAGATG 255
QY 247 GCAGCCAACTTCTCTGTGCGAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAATTT 306
Db 256 CTCCTGGACCAAGTATGAGGTGAATTAAGGTCTCAGCCCTGAGTGTGGATGCAAGATG 315
QY 307 GGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTGTGGGACTCATGCTTATGAC 366
Db 316 TTCCTCTATCTGACACTAGCAGGTTGGGAATTTTCTTCTTACGCCACCATGCGCTATGAC 375
QY 367 CACTAGTGGCGGTTAGCCACCACTTCATCTATCCCATCTCTCATGATGATCAGAGGCTGTG 426
Db 376 CGCTAGTGGCGGCTTGCATCTCTCCGTTTACCTGTCTCTCATGATGATGAGGCTGTG 435
QY 427 CTCAGATTACTGGGAGCTCTCTGGGCTTTGGGATTAATAGATGAGTGAATGATGAGTG 486
Db 436 CTTCCTGCGGAGGCTGTGTTGTTCTCTGAGTGAAGGCTTCTCATGCTCACTCCC 495
```

```

QY 487 GCACCATGGGCTTACCTTACTGTGGCTCAAGAGCGGTGATCACTTTTCTGTAGGTA 546
Db 496 ATGACCATGAGCTTCCCTTCTGTGAGATCCTGGAGATTCATCAATTTCTGTGAAGTC 555
QY 547 CAAGCTTATGAAGCTGGCGCTGTGACAGACACTTCCCTTTTGACACCCCTCTTTGCT 606
Db 556 CCTGCTGTAAAGATCTGTCTGTGTCAGACACCTCACTCTATGAGACCCCTCATAGTACTA 615
QY 607 TGTGTGTCTTCAATGCTTCTCTTCCCTTCTTCCATCATGATGAGGCTCTTATGCTTGCATC 666
Db 616 TGTGTGTCTTCAATGCTTCTCTTCCCTTCTTCCATCATGATGAGGCTCTTATGCTTGCATC 675
QY 667 CTAGGGGCTGTGCTCGAATACGCTCTGCTCAGGCGCTGAAAAAGCCCTGGCCACCTGC 726
Db 676 CTCTCTACCGTCCACAGGATGAACCTCAGCAGAGGCGCGAAAAAGGCCCTTGGCACCTGC 735
QY 727 TCCTCCACCTAACAGCTGTACACCTCTTCTATGAGGCGACAGCATGTTCAATGATCCTGAGG 786
Db 736 TCCTCCACCTGACTGTGTGTCATCTCTTCTATGAGGCGCTGCGGCTTACACCTCATGCTC 795
QY 787 CTTAGGCGCTACCGGCGCCCTTACCATGACAGGTGGCGCTTCTTACACAGTCTT 846
Db 796 CCAGCTCTTACACACCCCTGAGAAAGACATGATGTATCTCTTATACATCCTC 855
QY 847 ACTCCATGCTGAACCCCTCATTTTACAGCTTGAAGGAATGGGAGGTGATGGGCGCACTG 906
Db 856 ACTCCGCTGTGAACCTTTAATCTATAGTCTTGAAGGAATGATGATGATGAGGCTCTG 915
QY 907 AGGAA 911
Db 916 AAGAA 920
```

RESULT 11

```

US-09-886-055-434
; Sequence 434, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 434
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-434
```

Query Match 36.3%; Score 343.8; DB 10; Length 1047;  
Best Local Similarity 62.1%; Pred. No. 4e-99;  
Matches 543; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

```

QY 37 TCTCTCTTGGGCACTTTTCCACAGACGACCTGACCTTGTCTCTCTGACATTT 96
Db 130 TTCATCTGTGGGACTCTTTCAGACAATCCAAACATCCAGCACTACTTGTGTGTCATTT 189
QY 97 ATGCTGTCTTTCACAGTGGCGCTGTGGGAATGCTCTCTCATCTTCTCATCTGACCTG 156
Db 190 TTTGTGTTTCTCTGATGGCGTGTGTTGGAATGCTGTCTCATCTTGTGATACACTGT 249
QY 157 GACGCTGACTTTCACACCCCATGATCTTCTCTCAGGCACTCTCCCTCATGAGACCTC 216
Db 250 GACGCCACCTTTCACACCCCATGATCTTCTCTCATGATGATGATGATGATGATG 309
QY 217 ATGTTGTTCTGTACATTTGTGCCAAAGATGCAACCACTTCTGTGTGCAAGAGTCC 276
```



```
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 480
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-480
```

```
Query Match      35.8%; Score 339.8; DB 10; Length 954;
Best Local Similarity 60.8%; Pred. No. 7.2e-98;
Matches 554; Conservative 0; Mismatches 357; Indels 0; Gaps 0;
```

```
QY 16 AACCACTCTACACAGATGGCTTCTCTCTGGGACATTTTCCACAGCCAGACTGAC 75
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 13 AATTATTCGGTATGCGGACTTAACTCTCTGGGTTGTTTCAGCAACGCCCGTTCCCG 72
QY 76 CTGTCTCTCTCTCTCTGAGTTATGGTCTTTCACAGTGGCCCTCTGTGGGAATGCTCTC 135
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 73 TGGCTTCTCTTGGCCCTCAATCTCTCTGCTTTTGGACCTCCATAGCCAGCAACGTGCTC 132
QY 136 CTCATCTTCTCTCATCTCACTGAGCGCTGAGCTTCACACCCCATGTAATCTCTCTCAGC 195
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 133 AAGATCAATCTCATCCACATAGACTCCGCCCTCCACACCCCATGTAATCTCTCTCAGC 192
QY 196 CAGCTCTCCCTCATGAGCTCATGTTGGTCTGTAACATTTGCCAAAGATGGACCCAAC 255
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 193 CAGCTCTCCCTCAGGAGCATCTGTATATTTCCACCATTGTGCCCCAAATGCTGTGAC 252
QY 256 TTCCTGTCTGCGAGGAAGTCCATCTCTTGTGGGCTGTGGCATACAAATGGCTTTT 315
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 253 CAGGTGATGAGCCAGAGAGCCATTTCTTTGCTGATGCACTGCCCAACTTCTCTAC 312
QY 316 GTCTCTCTTGGGATGAGGGGCTCTTGTGAGCTCATGGCTTATGACCACTACGTG 375
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 313 TTGACCTTAGCAGGGGCTGAGTTCTTCTCTAGACTCATGTCTATGATCGCTACGTA 372
QY 376 GCCGTAGCCACCCACTTCACTATCCATCTCATGAATCAGAGGCTCTCTCCAGATT 435
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 373 GCCATCTGCAACCCCTGCACTATCTCTCTCATGAGCCGCAAGATCTGCTGTGATT 432
QY 436 ACTGGAGCTCTGCGCTTTGGGATATAGATGAGTGAATTCAGATGGTGAGCCATG 495
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 433 GTGGCGGAGCTGTGGGAGGGGTCTATGATGTTCTTGTCTACCCCGTCAACCATG 492
QY 496 GCGTTACTTACTGTGGCTCAAGAGCGTGGATCACTTTTCTGTGAGGTACAAAGCTTA 555
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 493 CAGTTCCTCTGTGCTCTCGGAGATCAACCACTTCTCTGAGGTGCTGCGCTT 552
QY 556 TTGAAGCTGGCTGTGAGACACTTCCCTTTTGGACACCTCTCTTGTGCTGTGTC 615
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 553 CTGAAGCTCTCTGACGAGACACATCAAGCTTACGAGACGCAATGATGTCTGTATT 612
QY 616 TTCAAGCTTCTCTCTCTCTCTCATCATGAGCTCTCTATGCTTGCATCTAGGGCT 675
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 613 ATGATGCTCTCATCTCTCTCTCTGATCTCGGGCTCTTACACAAGAAATTCATTA 672
QY 676 GTGCTCGAATAGCTGTGCTCAGGCGTGAAGAAAGCCCTGCGCACCTCTCTCCAC 735
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 673 GTTATAGGATGAGCGAGGAGAGGGGAAAGGCTGTGGCACCTGCTCTCAAC 732
QY 736 CTAAAGCTGTACACCTCTCTATAGGGGAGCCATGTTATGATGAGGCTAGGGGC 795
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 733 ATGCTGTGTGAGCTCTCTATAGGGCTGCGCATGTACATACAGCTGCTCATCT 792
QY 796 TACCGGCGCTAGCCATGACAAGGTGCTCTATCTTACACAGTCTTACTCCATG 855
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```
Db 793 TACCACACCCCTGAGCAGACAAAGCTGTATCTGCTTCTACACCATCTTACTCCATG 852
QY 856 CTGAACCCCTCATTTTACAGCTTGAGAAATGGAGGTGATGGGGCACTGAGAAGGG 915
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 853 CTCAATCCACTCATTTACAGCCTTAGAACAAGATGTCAAGGGGCCCTACAGAAGTT 912
QY 916 CTGACCGCTG 926
   |||||  |||||
Db 913 GTGGGAGGTG 923
```

```
RESULT 14
US-09-886-055-428
; Sequence 428, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; PRIOR FILING DATE: 2001-06-22
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 428
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-428
```

```
Query Match      34.8%; Score 330.2; DB 10; Length 1017;
Best Local Similarity 60.8%; Pred. No. 8.5e-95;
Matches 539; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
```

```
QY 40 TTCTCTTGGGCACTTTTCCACAGCCAGACTGACCTGTCTCTCTGACAGTATG 99
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 100 TTATCTTCTGTGGTTGTTTCAGCAACGCCGCTTTCCTGCGCTTCTTGGCCCTCATCTC 159
QY 100 GTGCTCTACAGTGCGCCCTCTGTGGAGTGTCTCTCATCTTCTCATCTACCTGAGC 159
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 160 CTGCTCTTGTGACCTTCATAGCCAGCAACGTGTCAAGATCATTTCTCATCCATAGAC 219
QY 160 GCTGACTTACACCCCATGTACTTCTCTCTCAGCCAGCTCTCCCTCATGAGCTCATG 219
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 220 TCCCGCTTCCACACCCCATGTACTTCTCTCAGCCAGCTCTCCCTCATGAGGACATCTG 279
QY 220 TTGCTCTTAACATTTGCCAAAGATGAGCAGCCAACTTCTCTGCGCAGGAAGTCCATC 279
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 280 TATATTTCCACCATTTGCCAAATGCTGTGACAGAGGTGATGAGCCAGAGACCAT 339
QY 280 TCCTTTGGGCTGTGCAATACAAATGCTTTTGTCTCTCTGTGAGATCTGAGGG 339
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 340 TCCTTTGTGATGACTGCCCAACACTTCTCTACTTGAACCTTAGCAGGGGCTGAGTTG 399
QY 340 CTCTTGTGGACTATGCTTATGACCACTACGTTGGCGTTAGCCACCACTTCACTAT 399
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 400 TTCTCTTAGGACTATGTCTGTGATGCTAGTACCATCTGCAACCTCTGCACTAT 459
QY 400 CCCATCTCATGAATCAGAGGCTGTCTCCAGATTAATGAGAGCTCTGCGCTTGGG 459
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 460 CCTGACCTCATGAGCCCAAGATCTGTGTTGATTTGGCGGACGCTGCGGAGGG 519
QY 460 ATAATAGATGAGTGAATTCAGATGTGAGCAGCCATGGGCTTACCTTACTGTGCTCAAG 519
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 520 TCTATGATGATTTCTGTCTCACCCCGTCAACATGCAATTCCTTGTGCTCTGCG 579
QY 520 AGCGTGAATCACTTTTCTGTGAGGTACAAGCTTTATGAAGCTGGCTGTGACACACT 579
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 580 GAGATCAACCACTTCTCTGCGAGGTGCTGCGCTTCTGAAGCTCTCTGCAAGGACACA 639
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

QY 580 TCCCTTTTGAACCCCTCCTCTTGTGCTGTGCTCTCATGCTTCTCCCTCTCC 639  
 Db 640 TCAGCCTACGAGACAGCCATGTATGTCTGTATATGATGCTCTCATCCCTTCTCT 699  
 QY 640 ATCATCATGCGCTTCTATGCTGTGACCTAGGGGCTGTGCTCCGAATACGCTGTGCTAG 699  
 Db 700 GTATCTCGGGCTTACACAGAATTTCTCATTACTGTTTATAGATGAGCGAGAG 759  
 QY 700 GCGTGAAGAAAGCCCTGGCCACCTGCTCTCCACCTAACAGCTGTACCCCTCTCTAT 759  
 Db 760 GGGAGGCGAAAGGCTGTGGCCACCTGCTCTCACACATGTGTGTGTGACGCTCTCTAT 819  
 QY 760 GGGCAGCCATGTTCATGTACCTGAGGCTTAGGGCTTACCGGGCCCTAGCCATGACAAG 819  
 Db 820 GGGGCTGCGCATGTACATACATACGCTGCTCATTTCTTACACACCCCTGAGCAGACAAA 879  
 QY 820 GTGCTCTATCTTCTACACAGCTTCTTCTCCATGCTGAACCCCTCATTTACAGCTTG 879  
 Db 880 GCTGTATCTGCTTCTACACACATCTCCTCCATGCTCATCATCATTTACAGCCTT 939  
 QY 880 AGGAATGGGAGGTGATGGGGGCTGAGGAAGGGGCTGAGCCGCTG 926  
 Db 940 AGGAACAAGATGTACGGGGGCTTACAGAAGGTTGTTGGAGGTG 986

RESULT 15

US-09-886-055-164  
 : Sequence 164, Application US/09886055  
 : Patent No. US20020132273A1  
 : GENERAL INFORMATION:  
 : APPLICANT: STRYER, LUBERT  
 : APPLICANT: ZOZULYA, SERGEY  
 : TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
 : FILE REFERENCE: 078003-0277150  
 : CURRENT APPLICATION NUMBER: US/09/886,055  
 : CURRENT FILING DATE: 2001-06-22  
 : PRIOR APPLICATION NUMBER: 60/213,812  
 : PRIOR FILING DATE: 2000-06-22  
 : NUMBER OF SEQ ID NOS: 522  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 164  
 : LENGTH: 972  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-886-055-164

Query Match 34.4%; Score 326.4; DB 10; Length 972;  
 Best Local Similarity 60.8%; Pred. No. 1.3e-93;  
 Matches 550; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

QY 15 GAACCACTCTACACAGATGGCTTCTCTCTGAGCATCTTTCCACAGCAGCACTGA 74  
 Db 12 GAATCAGACCTCTCTGAGACATCTCTGAGGGGCTTGTGATGACTCCCTTACCCA 71  
 QY 75 CTTGTCTCTTCTCTGAGTATGCTGTCTTCAAGTGCCCTCTGTGGAAATGTCT 134  
 Db 72 CTTTCTCTTCTCTGAGCAATGCTGTCTTCTTATTTGCGGTGAGTGCAACACCT 131  
 QY 135 CTTATCTCTCTCATCTACCTGAGCGCTGACCTTCAACCCCATGTACTTCTTCTCAG 194  
 Db 132 CACCATCTCTCTCATCTGATGATCCACAGCTTCAACCAATGTATTTCTGCTCAG 191  
 QY 195 CCAGCTCTCCCTCATGAGCTCATGTGTGTATTAACATTTGTGCAAGATGGCAAGCAA 254  
 Db 192 CCAGCTCTCCCTCATGAGCTCATGTGTGTATTAACATTTGTGCAAGATGGCAAGCAA 251  
 QY 255 CTTCTCTCTCTGAGGAAGTCAATCTCTTGTGGGCTGTGCAATTAACAATTTGCTTTT 314  
 Db 252 CTACCTATCTGCAAGAAATCTATCTCTTGTGGGCTGTGCAACCCAGACATCTCTCTA 311  
 QY 315 TGTCTCTCTTGTGGGATCTGAGGGGCTTGTGCTGGACTCATGGCTTATGACCACCTAGCT 374

Db 312 TTTGTCTAGTGTGCTGAATGTTTCTCTTAGCTGTATGCTATGACCGCTATGT 371  
 QY 375 GGGCGTTAGCCACCCACTTCACTATCCCATCTCATGATGACAGGGCTCTGTCTCAGAT 434  
 Db 372 TGGCATCTGTATCCACTGCGCTATGCTGTGCTCATGACAAAGAGTGGGACTGATGAT 431  
 QY 435 TACTGGAGCTCTGTGGGCTTTGGGATAATAGATGAGTGTATGATGATGATGATGAT 494  
 Db 432 GGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491  
 QY 495 GGGGTTACTTACTGTGCTGCTCAAGAGGCTGATGATGATGATGATGATGATGATGAT 554  
 Db 492 GCACTTCCCTTCTGTGGGCTGCGGAAAGTCTACCACTTCTACTGTGAGTTCCAGCTGT 551  
 QY 555 ATGAAGCTGGGCTGTGACAGACTTCCCTTTTGAACACCCCTCTTGTGCTGTGTGT 614  
 Db 552 TGTGAAGTTGTATGTGGGACATCACTGTGTATGAGACACAGATGTATCATGACAGCAT 611  
 QY 615 CTTCATGCTTCT 674  
 Db 612 TCTC---CTCCT 668  
 QY 675 TGTCTCTCGAATACGCTCTGTCTGACAGGCTGGAAGAAAGCCCTGCTCTCTCTCT 734  
 Db 669 TGTCTCTCAGATGGGCTCATCTGAGGAGCAAGAGAAATGCTTGGCCACTTGTGGCTCCA 728  
 QY 735 CTTACAGCTGTACCCCTCTTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 794  
 Db 729 CTTACAGGCTGTCT 788  
 QY 795 CTACCGGGGCTTACCATGACAGAGTGGGCTTATCTTCTACACAGTCTTACTCCAT 854  
 Db 789 CCAGTGCACCTCTATGACAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 848  
 QY 855 GCTGAACCCCTCATTTTACAGCTTGAAGATGGGAGGTGATGGGGGCTGAGGAAGG 914  
 Db 849 ATTGAATCTCTGATTTATATCTCTCCGGAATAAGATGTAGCTTAAGGCTCTGAGAGAGT 908  
 QY 915 GCTG 918  
 Db 909 GCTG 912

Search completed: May 9, 2003, 07:41:12  
 Job time : 1770 secs